Political Information, Elections and Public Policy

By

Valentino Larcinese

The London School of Economics and Political Science
Department of Economics

Thesis submitted to the University of London in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Valentino Larcinese

Political Information, Elections and Public Policy

Abstract

This thesis contributes to the study of the role of information in elections and public policy formation. Its main focus is on information acquisition and voting behaviour. Chapter 1 discusses the motivation of this research and presents a survey of related literature. Chapter 2 focuses on electoral turnout, Chapter 3 on public policy, and Chapter 4 on mass media.

Chapter 2 studies the impact of information on electoral turnout. Since incentives to be informed are correlated with other incentives to participate in public life, a model of information acquisition and turnout is introduced to isolate potential instrumental variables and try to establish a causal relation. Results are tested on the 1997 General Election in Britain. It is shown that information, as well as ideology, matters for turnout. It also contributes to explain the systematic correlation of turnout with variables like education and income. Voters' knowledge of candidates and of other political issues is also substantially influenced by mass media.

Chapter 3 presents a model that links the distribution of political knowledge with redistributive policies. It argues that voters can have private incentives to be informed about politics and that such incentives are correlated with income. Therefore redistribution will be systematically lower than what the median voter theorem predicts. Moreover, more inequality does not necessarily lead to an increase in redistribution and constitutional restrictions might have unintended consequences.

In Chapter 4 it is argued that instrumentally motivated voters should increase their demand for information when elections are close. In supplying news, mass media should take into account information demand, as well as the value of customers to advertisers and the cost of reaching marginal readers. Information supply should therefore be larger in electoral constituencies where the contest is expected to be closer, the population is on average more valuable for advertisers, and the population density is higher. These conclusions are then tested with good results on data from the 1997 General Election in Britain.

A Maria, Pasquale ed Evelina

Acknowledgments

I am grateful to all those who devoted their time to listen, discuss, criticise, and make me understand things when possible. My greatest intellectual debt is to my advisor Tim Besley who stimulated my interest in political economy, has provided invaluable guidance on my research and on its methodology, and has represented a constant source of enthusiasm in these years. A number of other people helped me in developing this work, with conversations, comments, and precious suggestions: I would like to thank Fabrizio Balassone, Oriana Bandiera, Richard Blundell, Andrea Brandolini, Michela Cella, Harold Clarke, Giacomo Corneo, Guillaume Frechette, Jean-Francois Laslier, Cesar Martinelli, Torsten Persson, Michele Piccione, Paolo Ramezzana, Imran Rasul, Leo Rizzo, Kevin Roberts, Howard Rosenthal, Paul Schweinzer, Cecilia Testa, Cristiana Vitale, Stan Winer, and Donald Wittman. I am also indebted with participants in seminars and meetings where I had the opportunity to present my work. My sister Evelina Larcinese helped me in the process of collecting data from newspapers, representing for a few months the best conceivable of research assistants. Beyond the limits of this work, I am very grateful to Maria Pia Monteduro, for the long and challenging conversations we had on this as well as many other subjects: they contributed a lot to the development of my ideas.

In this time STICERD has been just the ideal place where to work, both for the intellectual climate and the informal and friendly atmosphere: I wish to express my gratitude to all the people there. I have also benefited from discussions with colleagues while working in Banca d'Italia and at the Government Department of LSE, and during a period spent visiting Universitat Autonoma in Barcelona.

A number of people have been patient with me during the completion of this thesis: in particular, the understanding and support of my friend Cristiana Vitale has been very important. The patience of Chai Lieven and Keith Dowding, in the LSE Government Department, is also very much appreciated.

It is sometimes the case that the process of getting somewhere is better than its results: in my case, while working on this dissertation, I enjoyed very much the time spent at LSE as a PhD student, thanks also to a number of friends and exceptional people that I had the luck to meet. I thank them all, the list could be very long and I would rather not take the chance of omitting anyone, but I would like in particular to express my gratitude to Imran Rasul, office mate for a long time, before the open plan era, and probably the person with whom I have shared most of this experience, in both its academic and less intellectual facets.

I also wish to express my gratitude to those who provided financial support for my studies: I gratefully acknowledge scholarships received from the European Commission, ESRC, and Consiglio Nazionale delle Ricerche.

Obviously, I am the sole responsible for limits and errors of this dissertation.

Contents

List of Tables

List of Figures

"Nulla si sa, tutto si immagina"
(Nothing is known, anything can be imagined)

[Federico Fellini, among others]

Chapter 1

Information, Elections, and Public Policy

1.1 Introduction

Recent years have witnessed an increasing attention in economic literature on the process that leads to public policy formation. There has been a growing consent on the view that to evaluate policies and their effectiveness it is necessary to open the black box that leads to policy choice.

An economy based on decentralised and voluntary exchange can only achieve Pareto-efficiency if some restrictive assumptions are satisfied. When such conditions are not met, i.e. market failures or incompleteness occur, then government intervention can help reaching more desirable outcomes. Traditional public economics assumes that policy-makers are benevolent welfare-maximizers. Under this assumption, public policy-making, and its outcomes, will only be constrained by the range of policy instruments and by information availability. Thus, perfect governments could reach the same outcome as perfect markets.

However, these institutions are hardly perfect in reality and it has become clear that an important purpose of economic science should be to better understand the functioning of both. After devoting much effort to studying market imperfections and their consequences, economists seem to have turned their attention to government imperfections, thus substantially modifying both positive and normative analysis of government action.

Opening the black box of politics is a complex task. Public policies are the outcome of a number of interactions between often conflicting agents. Most of this analysis has been focussed on the study of liberal democracies: here citizens express their preferences electing their representatives, politicians compete for consensus, pressure groups and organized private interests try to affect public policy through their influence on public officials or public opinion, and bureaucrats and other non-elected officials are often responsible for policy implementation and have the opportunity to leave their mark on final outcomes. Research by political scientists and, more recently, by economists, has tried to understand the complex interactions between the agents involved in political processes, and the role of the rules of the game in determining policies.

In this picture, the election of public officials assumes a central and non-substitutable role: it is in the intention of any implicit or explicit democratic constitution to leave the ultimate power to citizens. Moreover, although many specific policies can probably be affected by organized interest groups better than by citizens, broad policy choices, concerning for example income taxation, pensions, health services, should ultimately rest on general consensus of, at least, a relative majority of the population. Thus, a central task has become understanding the functioning of electoral systems and their capability of representing and aggregating citizens' preferences.

Do elections actually serve the purpose of linking citizens' preferences and elected officials behaviour? The answer to this question depends on the functioning of electoral institutions. As proved by Arrow (1951), it is impossible to aggregate a given set of individual preferences into a collective decision using a decision-rule that satisfies some minimal consistency and procedural requirements. Thus, institutions matter. The same set of preferences can be aggregated in a number of different ways and the final outcome will inevitably depend on the rules of the game.

The formal literature on elections has therefore devoted most efforts to understand the impact on policy of both preferences and institutions, placing restrictions on either or both in order to gain predictive power. Most of this literature seems to take for granted that citizens have enough knowledge of political institutions, electoral platforms and the consequences of given policies, to be able to pursue their interest in the electoral process. The potential role of information is, in some cases, simply ignored. This is at odds with the fact that economists have long ago recognized the key role of information in decision-

making. The link between preferences and behavior is mediated by beliefs about the external world and about other players. In the same way, the link between preferences and public policies is mediated not only by institutions, but also by the knowledge people have of political and economic matters.

In reality, citizens often appear to have little competence in political matters. Since the origins of democratic forms of governance, political theorists have often supposed that citizens lack adequate knowledge to govern themselves wisely. Plato, for example, claimed that in the "good republic" ultimate power should only be given to "philosophers", people with the competence to understand where the public good lies and how to reach it. Aristotle was also skeptical about individuals' civic capabilities, but less so about democracy in itself: "There is this to be said for the Many. Each of them by himself may not be of a good quality; but when they come together it is possible that they may surpass - collectively and as a body - the quality of the few best"[1]. As it will be clearer later, the question of whether a group of poorly informed individuals can take good collective decisions is still one of the most intriguing and debated in political science.

Skepticism about the civic virtues of citizens have survived until very recently. Restrictions to civil rights are common to all democracies, where for example people are not entitled to vote before a certain age. Less justifiably, restrictions have been in place until quite recently against women, some racial groups, or people with low income or educational attainment. It is only during the XX century that universal suffrage has become accepted as one of the main features of proper democratic governance.

While very few today would express doubts that the ultimate power in a community should be with the whole body of its citizens, the question of the role of political information in elections has become more crucial than ever. An extensive body of empirical research conducted in the last fifty years confirms that the general public is often poorly informed about political issues. Moreover, there are systematic correlations between being economically endowed and being politically informed. Also, the level of participation in democratic life shows a strong correlation with characteristics such as income or education.

Nevertheless, the role of information in the political market is the object of a very open debate. On the one side, it is possible to argue that information is essential to good

[1] Aristotle (1946), p.123.

decision-making and that therefore the levels and distribution of political knowledge will affect both the accountability of public officials and the distribution of political influence. On the other side, little information could be irrelevant or almost so. This has been argued along two lines: the "behavior irrelevance" advocates claim that, for a variety of reasons, uninformed voters manage to vote as if they were informed; the "outcome irrelevance" champions, instead, argue that uninformed voters do behave differently, but that the final electoral outcome is the same that a perfectly informed electorate would choose. Each of these views clearly carries its own policy and normative implications[2].

The main purpose of this thesis is to contribute to the understanding of the determinants and consequences of citizens' knowledge of political issues. Our methodological starting point is rational choice theory. Thus, our interpretation of observed levels of political knowledge is that it is the outcome of individual decision-making within a given choice set. However, knowledge or ignorance are purely the consequence of individual choice when people face the same opportunities to learn. In this sense, preferences matter. In reality, individuals face different choice sets for a number of reasons. This can be due to individual as well as environmental circumstances. At the individual level, factors like income, time, and personal connections, provide the means to become informed. Furthermore, some personal circumstances can influence the capability to acquire, process and retain information: innate skills, formal education, experience and so on. Some characteristics of the environment also determine how much and what people know. Factors like news availability, independence of the mass media, competition among different information sources, all provide more opportunities to get more and better information. Moreover, the relevance, the closeness or even just the symbolic importance of an election (like for the recent second round presidential election in France) provide environmental motivations to become better informed.

This thesis will present a model of information acquisition that determines citizens' political knowledge as an equilibrium choice that depends on preferences, prior beliefs, and external and internal constraints. We will show that political knowledge matters for public policy and that the asymmetry in incentives to be informed faced by different citizens is especially important for redistributive policies. We will also consider the role of information supply by mass media and the potential distortions that they introduce.

[2]The contributions to this debate will receive the due attention later in this chapter.

Whenever possible, theoretical predictions are tested using data from the 1997 General Election in Britain (England, Wales and Scotland). A more detailed overview of the thesis is given at the end of this chapter.

We will now try to put this work in the appropriate context by discussing some of the related literature. Like for all reviews, some choices are in order. Given the exponential growth of this field in the last twenty years, we will discuss only few economic models of elections, and the main results they deliver, with special attention for redistributive policies. We will then turn to voters' knowledge of political matters. Economists' research in this case is still quite limited, whereas the issue has long been debated in political science and communication studies, where a vast empirical literature has established a large number of stylized facts. Finally we will discuss the main themes concerning the consequences of citizens' political knowledge or ignorance for voting behaviour and public policy-making. Here the debate seems very open and very different views have been expressed along the centuries, as well as in the last ten years.

1.2 Elections and public policy-making

Elections affect public policies. This should be true in any country that can be defined as democratic. However, the causal links between the two can take several different paths, with different implications for public policy. The first step is to recognize that public officials have their own motivations and policy preferences. Schumpeter's *Capitalism, Socialism and Democracy* (1942) is among the first works in economic literature to criticise the naive view of democracy of early welfare economics. For Schumpeter, policies are not the expression of the common will and politicians do not necessarily seek the common good. Instead "the democratic method is that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote".

However, the truly devastating result for traditional views of democracy comes a few years later in Arrow's *Social Choice and Individual Values* (1951), which explores the logical possibility of aggregating individuals' preferences into a common will and reaches the conclusion that "if we exclude the possibility of interpersonal comparison of utility, then the only methods of passing from individual tastes to social preferences which will

be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial". Thus, Arrow leaves us with a big question mark in place of commonly used expressions like "common good" or "common will". Hence, not only public officials do not necessarily represent the common will: the common will in its most abstract and absolute meaning does not even exist.

1.2.1 The Downsian model

This impossibility result stimulated research in the area of social choice and many attempts have been made to overcome the problem, mainly by weakening one or the other of Arrow's requirements. Among these several paths of study the so-called *public choice* approach is characterized by a quite pragmatic view of the problem: restrictions are imposed on the institutional setting (majority rule), on the policy space (single issue), and on preferences (single-peakness); this leads to formulate models with sharp predictions. Anthony Downs' *An Economic Theory of Democracy* is probably the first attempt to analyze real world's elections using the tools provided by economic theory. This pioneering work in 1957 starts what could be called an economic theory of politics. The notion of *homo oeconomicus*, rationally acting for her selfish ends, is introduced in the world of politics: politicians are defined as acting "solely in order to attain the income, prestige, and power which come from being in office. (...) Upon this reasoning rests the fundamental hypothesis of our model: parties formulate policies in order to win elections, rather than win elections in order to formulate policies"[3].

The striking feature of the Downsian model is that it deals with politics as economists deal with markets, seeing the political process as an exchange: consent and policies are exchanged in a specific institutional setting in the same way as commodities are exchanged on markets. Downs' work was able to provide an explanation, by using a unified theory, of political phenomena like party convergence, voter turnout, or disinformation on public issues. Since these issues had been puzzling political scientists for a long time, Downs's theory could either be accepted with enthusiasm or harshly criticized, but surely not ignored.

A central result of the Downsian model is that, if all citizens' preferences are single-

[3] Downs (1957).

peaked on a single dimension, then the median ideal preference is a Condorcet winner[4]. Two main requirements are crucial: 1) a Condorcet winner exists; 2) there is competition among parties to reach it. In this case two competing parties, able to pre-commit to their proposed platforms, will converge on the policy preferred by the median voter. Thus, although under quite restrictive assumptions, Downs delivers a clear prediction about the impact of electoral competition on public policy.

It is not our purpose here to discuss all the limitations of the Downsian model. It cannot be overlooked that, probably because of easiness of interpretation and of its clear-cut predictions, Downs' model has been the most influential work on electoral competition.

1.2.2 Redistributive policies

Applications of the median voter theorem have been used to explain virtually every dimension of public policy-making. A vast body of research has used the Downsian model to explain income redistribution through fiscal policy: when the poor are a majority and gross income is fixed, then we should expect "slavery of the rich" (Foley, 1967), i.e. the poor will redistribute all available income in their favour by voting the highest possible tax rate. Romer (1975) and Roberts (1977) generalize this model introducing the possibility that rich agents react to taxation by reducing labour supply. This implies that governments can only use distortionary taxation. Roberts (1977) shows that if voters redistributive preferences are monotonically related to their productivity then a Condorcet winner exists and it is the redistributive tax preferred by the voter with median productivity[5]. In Roberts' model, equilibrium income redistribution (linear tax rate and lump sum benefit with balanced budget) depends on the distance between the median voters' income and the mean income in society. Since real income distributions are right skewed, this implies positive tax rates in equilibrium. At the same time, the median voter is aware that there is a negative link between the tax rate and total gross income and finds optimal to vote for a tax rate below 100%. Roberts' analysis has been used by Meltzer and Richard (1981) to provide an explanation of the rapid growth of the public sector

[4]See also Hotelling (1929) and Black (1948). A Condorcet winner is defined as an alternative that can beat any other alternative in the policy space in pairwise comparison and by using majority voting.

[5]Roberts calls this monotonicity condition "hierarchical adherence". Recently, Gans and Smart (1996) have shown that Robert's condition is substantially equivalent to the Spence-Mirrlees condition of indifference curves single-crossing.

observed for most of the last century. They use the idea, first developed by de Tocqeville (1835), that the size of a government, measured by tax revenue and expenditure, depends essentially on the spread of the franchise and the distribution of wealth: the "extension of the franchise to include more voters below mean income increase votes for redistribution and, thus, increase this measure of the size of government".

In a static framework like the one considered by Roberts, the productive input is usually identified with labour supply. However, it is possible that the disincentives introduced by redistribution lasts over time by affecting investment decisions and ultimately growth rates. This possibility has been scrutinized by Bertola (1993) and Alesina and Rodrick (1994) within the Downsian framework. They consider a population in which agents are alike in all respects except for their initial ownership shares in the economy's aggregate stocks of capital and labour. Taxes are levied on capital income and revenue is used either for pure redistribution (Bertola) or to enhance the productive capability of the economic system (Alesina and Rodrick). These papers show that the poorer the median voter's income relative to the mean income, the higher the capital tax rate and the lower the growth rate. Persson and Tabellini (1994) reach similar conclusions in an overlapping generations model in which personal (and not factor) income distribution is considered.

These papers share a common limit in that they assume myopic behavior by agents. When voting or investing, agents are not taking into account the effects of their choices on future political equilibria[6]. The problem is that, by requiring the existence of a Condorcet winner, the Downsian model finds limited application in dynamic frameworks, where the probability of finding a Condorcet winner is much reduced.

The Downsian analysis of redistributive policies delivers two sets of testable predictions: 1) input supply decreases when government transfers and distortionary taxes increase; 2) government transfers increase with the distance between the mean and the median income.

[6] As Alesina and Rodrick themselves recognize "in our model the distribution of assets is predetermined and remains constant. In reality growth itself affects income distribution. The serious technical problem introduced in this case is that when income distribution varies over time, as a function of growth, one cannot look at each voting decision in isolation. Voting decisions in each period affect growth in subsequent periods, which, in turn affects distribution and future voting decisions. Thus, the outcome of future social choices depends on the voting decisions taken today. Therefore, when voting today, rational voters should internalize the dynamic problem of social choice". See Alesina and Rodrick (1994).

Persson and Tabellini (1994) and Alesina and Rodrick (1994) test a reduced form relation between inequality and growth, and find this relationship significantly negative. The use of a reduced form, however, casts some doubts about the intermediate steps, namely the positive relationship between inequality and redistribution and the negative relationship between redistribution and investment. Indeed, Perotti (1994), estimates a cross-country structural form of this class of models and finds that the "results concerning the two mechanisms examined here are conspicuously inconsistent (indeed, opposite to) the theory and conventional wisdom". The first result is that government transfers have a positive and significant effect on investment. The second is that a "higher share of the third quintile increases the share of government transfers in GDP". Of course for many countries considered in Perotti's analysis the process that leads to policy-making cannot properly be defined as democratic, so we cannot expect the median voter theorem to hold. However, what is surprising is that the positive effect of the third quintile share on transfers is even stronger in democracies. The t-statistics of the variables involved, even though not brilliant, can be considered enough to "cast some doubts on the empirical validity of the endogenous fiscal policy explanation of the relation between income and investment".

An analysis using panel data for OECD countries from 1960 to 1981 by Lindert (1996) finds that "wider inequality in pre-fisc incomes significantly reduces total government spending as a share of GDP (...). The anti-spending effect of inequality is spread across all (...) spending categories except unemployment compensation, which tends to be the smallest of these spending categories".

There is a different theory linking income inequality and public spending that, according to Lindert's findings, receives better support from the data. This is the social-affinity theory which predicts more redistribution the closer middle-income voters feel they are to the poor and the further they feel they are from the rich[7]. This theory shares similarities with a model proposed by St.Paul (1994) where it is shown that more inequality is even compatible with less distance between the median and the average income. If the lowest

[7]Even though apparently this concept is similar to that of Roberts, it involves a different specification of the explicative variables. Here we don't consider the mean/median ratio, but two gap variables: income of the top quintile over income of middle quintile (upper income gap) and middle over bottom quintiles (lower income gap). As stressed by Lindert "the social affinity hypothesis could, but need not, be narrowed to predict a positive effect of income skewness (...) on progressive social spending. It makes no prediction about the effect of inequality on social spending".

income levels below the median become poorer this does not affect the median income, though reducing mean income and, therefore, the desire for redistribution of the decisive voter. This dynamics leads to a situation of social exclusion, in which the poor may become poorer, but this does not lead the middle classes to support more redistributive programs.

Lindert's analysis shows that the social-affinity hypothesis receives support from the data. "The coefficient of the upper gap is positive, and that of the lower gap is negative, for clearly progressive redistributions". However, this result does not hold for pensions and health. Hence "all of the results would be consistent with the social affinity hypothesis if the progressivity ranking of the different clusters of tax-based social spending were, and were perceived to be, [total-social, welfare, unemployment, and education]>[pensions and health]. Yet it is not clear that education belongs in that more progressive category, nor is it clear that the pension and health programs are much less progressive. With this disclaimer, the overall pattern of social spending results appears to support the social-affinity hypothesis".

However by no means conclusive, it is clear from current empirical analysis that the Downsian models of income redistribution receive little support from empirical analysis. This could be explained in many ways; for example, lobbies' influence on policy-making can make politicians more independent of public opinion than what postulated by electoral models. Another possibility is that the rich have more power and influence on policy[8]. This solution, however, is just assumed and the type of power the richer groups have is not well specified. In the third chapter of this thesis we will provide an information-based rationale for this asymmetry of power in elections.

1.2.3 Beyond Downs

The main problem with the Downsian model is that even small complications in the basic framework can destroy its predictive power, as the existence of a Condorcet winner is ensured only under restrictive conditions. For example, distortionary taxation in Roberts' model is a technically necessary assumption for a Condorcet winner to be found. Although assuming that taxes are distortionary is not particularly disturbing, things get more serious if we want to study multiple policy dimensions: in this case very strong restrictions

[8]This possibility is for example considered in Benabou (2000) and Saint Paul and Verdier (1997).

are required on the distribution of preferences in order to have an equilibrium (Plott, 1967; Grandmont, 1978). Nevertheless, many problems of interest for an economist involve such multiplicity: for example, the case of a government that decides on both public good provision and progressive taxation. Dynamic settings pose further problems for the existence of a Condorcet winner and restrictions on voters' rationality are in this case required to ensure the existence of an equilibrium.

Several solutions have been proposed to overcome these limits. Voting on different issues one by one is a possibility. This makes possible to find a "structure induced equilibrium" when there is no Condorcet winner. Kramer (1972) proves that, with separable preferences, the issue-by-issue stable point is the issue-by-issue median preference. However, if preferences are not separable, the final equilibrium will inevitably depend on the sequence in which the different issues are voted. Even the existence or not of an equilibrium may depend on the voting agenda. Thus, whoever sets the agenda has a disproportionate influence on the final collective decision. Miller (1980) shows that this influence can, to a certain extent, be counterbalanced by voters' sophistication. Independently of the agenda, voters can always manage to choose options within a specific subset of the policy space (uncovered set): in a two party competition this forces candidates to propose platforms within the uncovered set (Shepsle and Weingast, 1984; McKelvey, 1986). This result provides a characterization of the complex interaction between rational behaviour and institutions.

Another way to look at the problem of the existence of a Condorcet winner is to focus on the discontinuities in the payoff functions of the parties. This discontinuity derives from the fact that small changes in the platform proposed can lead a loser candidate to become a winner and vice-versa. Transforming this discontinuous deterministic payoff function into a continuous probability of victory can help in finding an equilibrium. This consideration led to the development of the so-called probabilistic voting model (Hinich, 1977; Coughlin and Nitzan, 1981), that has found applications in a number of works[9].

Nevertheless, there are important features of real elections that are not captured by

[9] Of course, the reason for introducing this class of models is not just a technical one. Calvert (1986) justifies the use of probabilistic voting models on the ground that assuming that "candidates cannot perfectly predict the response of the electorate to their platforms is appealing for its realism". To use the words of Coughlin (1992), "because of their importance to candidates' decisions, the candidates' beliefs about how their choices relate to the voters' choices provide a natural dividing line for the economic models of elections that have been developed".

any of the models presented above. The issue-by-issue voting, for example, is mostly typical of legislative or other bodies, but in political elections citizens have to choose between candidates, and do not have the possibility of unbundling different issues. This brings us to an assumption which is crucial both for the Downsian and the probabilistic voting models: that candidates can credibly pre-commit to their proposed platforms. This is a plausible assumption only if candidates are merely office-seeking. In this case they would be indifferent to any specific policy and may well implement the promised one (provided such implementation has no additional costs as compared to other possible policies). However, political actors, as well as the voters, may care about policies. A different class of models explores this alternative interpretation of candidates' motivation. In these models, candidates are assumed to be partisan: they care about winning in order to be able to implement their preferred policy, rather than the reverse. Nevertheless, in a majoritarian system preferences can be translated into policies only in case of victory, which gives candidates an incentive to move toward the median voter. This tension between policy preferences and the incentive to win provides new insights in the understanding of electoral competition[10].

The problem with this approach is that the political platform proposed by each candidate is not time-consistent. As stressed by Alesina (1988), the incentives faced by the candidate change after the election: there will be no need to target the median voter and the policy preferred by the candidate rather than the one proposed during the campaign will be implemented. If voters anticipate this process, there is no way for a candidate to make a credible commitment to any policy different from her preferred one. This credibility problem may lead to non-convergence. Repeated interactions can help to overcome this credibility problem through reputation. This also provides an explanation for the presence of long-lived actors like parties in the political arena. As in all applications of the folk theorem, for reputation to be effective the discount rate must be sufficiently high, in the sense that parties should care enough about the future. Parties are faced by a trade

[10] Calvert (1985) shows that if the median voter preference is known and candidates' policy preferences are on opposite sides with respect to the median voter, then there is perfect convergence to the median voter's preferred policy even with completely partisan candidates. The reason for this is that a candidate can win elections being slightly closer to the median voter than its opponent. Then, for any given platform of the opponent, the policy outcome will be closer to the preferred one when closer than the opponent to the median voter. This eventually leads both candidates to target the median voter: candidates' motivation changes but the political implication is the same as in the Downsian model.

off between acting unconstrained and enjoying immediately all the benefits of victory or instead "spreading" the benefits of victory over time. Enjoying a high initial reputation among the public gives candidates (Alesina, 1988) or incumbent public officials (Coate and Morris, 1995) the possibility to have more immediate returns. This indirectly points to the role of information as a disciplining device.

The need to introduce more features of real electoral processes is also recognized in the so-called citizen-candidates approach. In two separate papers Osborne and Slivinsky (1996) and Besley and Coate (1997) propose a new model of electoral competition which explicitly recognizes the fact that most policy decisions are taken in a context of representative democracy[11]. One important feature of this model is that candidates run for office with their own preferences about policies: this means that they will not be able to pre-commit to anything else than their preferred policy outcome. The role of the platform announcement in the Downsian model is thus replaced by an entry stage in which each citizen may enter the political competition at a given cost.

In the citizen-candidate model we can expect to find equilibria in many situations in which there is no Condorcet winner and therefore the Downsian model would not provide any prediction. On the contrary, this model gives "too many" equilibria. This is because it gives only a minimal institutional structure to the electoral process, unveiling the possibility of having many potential equilibria, where people's beliefs and further institutional constraints are then essential to understand where the electoral process leads in terms of policies. "For those who would like a clean empirical prediction, our multiple equilibria will raise a sense of dissatisfaction. However, this findings squares with the more familiar problem of game theoretic models: that rationality alone does not typically pin down equilibrium with complete precision (...). This suggests the need to understand better the role of political institutions as coordinating devices, giving some greater determinacy to equilibrium outcomes".

Being able to derive predictions from multidimensional models has relevant implications for the study of redistributive policy. Roemer (1999), for example, explains progressive taxation (without complete expropriation) in a model with exogenous income. This is

[11] In the words of Besley and Coate "the primitives of the approach are the citizens of a polity, their policy alternatives, and a constitution which specifies the rules of the political process. (...) No pre-existing political actors are assumed, and no restrictions are made on the number or type of policy issues to be decided. Political outcomes are thus derived directly from the underlying tastes and policy technology".

done by modelling internal party conflict among groups with different objective functions. Besley and Coate (1997 and 1998) show how the interaction between redistribution and other policy dimensions can lead to outcomes that are substantially different from the Downsian predictions. Using the citizen-candidate framework, Besley and Coate (1998) also analise dynamic models of redistribution without posing restrictions on the rationality of voters. In this context, they show that the lack of commitment to redistribute can generate inefficiencies, as Kaldor-Pareto improving policies could not receive the support of the majority.

1.3 Perfect information?

One characteristic most models of electoral competition have in common is that they assume that citizens are perfectly informed[12]. They are assumed to be aware either of announced platforms or of candidates' tastes and competence, depending on which model we consider. Moreover, voters are capable of perfectly predicting the effects of different policies, a capability any economist would be very happy to share with them! In the probabilistic voting model candidates' uncertainty on people' preferences is recognized but not voters' uncertainty on platforms or candidates characteristics. In reality, information seems to play a key role in elections. A clear sign of this is that politicians are extremely concerned with the amount and the quality of the coverage they get from mass media; they often seem to struggle for media attention using all available means.

Although the standard Downsian model assumes that everyone is perfectly informed and votes, Downs himself was aware that putting together rational behaviour and political participation is a challenging task. In a sizable electorate "the returns from voting are usually so low that even small costs may cause many voters to abstain"[13]. This tendency should be reinforced by parties' convergence. If the probability to be a pivotal voter is extremely low (for example the electorate is very large) then even small voting costs should induce massive abstention. Even if observed abstention sometimes may be very high[14], it is nevertheless difficult to explain the behaviour of millions of voters in this way.

[12] For exceptions see Ledyard (1984) and Palfrey & Rosenthal (1985).

[13] Downs (1957).

[14] Some evidence has also been found that in "close" elections (for example when polls are very uncertain, implying an higher probability to be pivotal) the turnout has been higher (Morton, 1991)

Another manifestation of political interest is to become informed about political issues. Also in this case, citizens have low incentives: "a rational man can become well informed for four reasons: 1) he may enjoy being well informed for its own sake, so that information as such provides him with utility; 2) he may believe the election is going to be so close that the probability of his casting the decisive vote is relatively high; 3) he may need information to influence the votes of others (...); 4) he may need information to influence the formation of government policy as a lobbyist. Nevertheless, since the odds are that no election will be close enough to render decisive the vote of any one person, or the votes of all those he can persuade to agree with him, the rational course of action for most citizens is to remain politically uninformed" (Downs, 1957).

That people are substantially ignorant on most political issues is supported by extensive empirical evidence, mainly conducted on the United States. The first systematic evidence on this came from the public opinion studies conducted between the 1940s and the 1960s. From their study on the 1952 and 1956 elections, Campbell *et al.* conclude that the electorate "knows little about what government has done (...) or what the parties propose to do. (...) The mass electorate is not able to appraise either its goals or the appropriateness of the means chosen to serve these goals" (Campbell *et al.*, 1960, 543). Converse (1964) found that only 10 per cent of the interviewed could define the meaning of words like "liberal" or "conservative". More recent studies do not change substantially the picture. According to Neuman (1986) "even the most vivid concepts of political life (...) are recognized by only a little over half the electorate".

Evidence on what voters know is not more encouraging. Delli Carpini and Keeter (1996) report that during the 1992 presidential campaign "86 per cent of the public knew that the Bushes' dog was named Millie, yet only 15 per cent knew that both presidential candidates supported the death penalty". Going beyond anecdotic evidence, they report that "only 13 per cent of the more than 2000 political questions examined could be answered correctly by 75 per cent or more of those asked, and only 41 per cent could be answered correctly by more than half the public". Political knowledge also shows a remarkable stability over time: "in spite of an unprecedent expansion in public education, a communication revolution that has shattered national and international boundaries, and the increasing relevance of national and international events and policies to the daily lives of Americans, citizens appear no more informed about politics" (Delli Carpini and Keeter,

1996).

However, the evidence about people's little knowledge of politics is not evidence *per se* of a lack of motivation (as Downs' theory would suggest). Key (1966) suggested that voters are as informed as the political context allows them to be, in the sense that often candidates are vague and not neatly distinguishable on specific issues. For example, a number of studies conducted during the 1960s and the 1970s show an increase in political knowledge compared to previous decades, probably related to the events occurred at the time and the consequent radicalisation of electoral platforms (Pomper, 1972; Nie, Verba, Petrocik, 1976).

When the context, as well as other factors, is taken into account, it is not clear if this evidence should be regarded as people having "too little" information or "too much". Indeed, according to Downs, we should expect the electorate to know even less than they seem to know. In general, learning requires motivation, ability, and opportunity. When we observe a given level of knowledge, still little can be said about its determinants: a limited knowledge of politics can be due to a lack of individual interest, or to poor ability, or rather to environmental circumstances that affect the opportunity to learn.

One very important component in this picture is the role of electoral campaigning and of the mass media. Research in this direction started in the period between the two World Wars, under a general presumption that mass communication was an extraordinarily powerful device: "it is no daring prophecy to say that the knowledge of how to create consent will alter every political calculation and modify every political premise" (Lippman, 1922). These theories go now under the name of "theories of mass propaganda": use of the media for political propaganda was quite common at the time both by authoritarian regimes and by the Allies during the war. However, the first systematic study conducted on survey data by a group of researchers at Columbia University seemed rather disappointing. Lazarsfeld, Berelson and Gaudet (1944) studied the Eire county during the 1940-44 American elections, finding little evidence in favour of the theories of mass propaganda: "The people who did most of the reading and listening not only read and heard most of their own partisan propaganda but were also most resistant to conversion because of strong predisposition. And the people most opened to conversion - the ones the campaign manager most wanted to reach - read and listened least" (Lazarsfeld *et al.* 1944). The influence of the Columbia school (also through a subsequent work

by Berelson, Lazarsfeld and McPhee, 1954) was such that since then the dominant view has been that campaigning and the media have only "minimal effects" on voters. Until quite recently, most studies continued to find little evidence of persuasion by mass media (Finkel, 1993).

A radical shift in communication studies has been induced by a new cognitive theory that goes under the name of "uses and gratifications"[15]. Rather then asking what are the effects of the media on people's opinions, this theory starts by asking why the people use the media in the first place. Only understanding individual motivations it will then be possible to recognize the possible effects. It is immediate that this theory should be of particular interest to rational choice theorists as it basically starts from individual preferences.

The shift in the focus of attention produced new empirical evidence that seemed in contrast with the minimal effects found by the Columbia school. In particular "the news can be expected to influence public opinion directly through three main avenues: enabling people to keep up with what is happening in the world and mobilizing them to vote (civic engagement), defining the priority of major political issues (agenda setting), and shaping people's political preferences (persuasion). In turn, these attitudes can be expected to influence reasoned voting choices" (Norris *et al.*, 1999). Iyengar and Kinder (1987) examine evidence from electoral campaigns and television news and conclude that their effects have not much to do with persuasion but rather with "commanding the public's attention (agenda-setting) and defining criteria underlying the public's judgement (priming)"[16]. Bartels (1993) shows how apparent "minimal effects" can be, at least partially, a consequence of measurement errors. Zaller (1992 and 1996) puts forward one further argument against the "minimal effects" evidence by arguing that tangible effects are only due to the "reception gap", the difference between the amount of information received about different candidates. According to Zaller, most studies were conducted on presidential elections, where the campaign is normally quite intense on both sides, with plenty of information on both candidates: this generates a minimal reception gap and therefore minimal effects, which is not the same as saying that the campaigns had no effect. In local elections, where the reception gap between incumbents and challengers is

[15] See Blumler and McQuail (1968).

[16] Bartels (1988), Zaller (1989), Popkin (1991), and Franklin (1991) find similar results.

normally much larger, the impact of the media appears instead sizeable.

It is useful at this stage to try to be more precise about what we intend for political knowledge. "Everything is politics" was a successful slogan in the seventies: this would imply that any knowledge is political knowledge. Even if our idea of electoral politics is that, luckily, not everything is politics, nevertheless the range of issues and facts that are relevant to politics is extremely large. How can we gauge political knowledge? What is relevant and what is not? Is there any risk to bias our assessment with our priorities? Delli Carpini and Keeter (1996), in presenting evidence on political knowledge of American voters, based their analysis on nearly 3700 questions collected in various surveys. They concluded that "researchers developing national or general political knowledge scales need not be overly concerned with the mix of specific topics covered by individual items. Scales made up of items tapping only knowledge of institutions and processes, substantive issues, or public figures are likely to serve as reasonable measures of the overarching construct"[17]. This is extremely important to us. The empirical analysis on information acquisition presented in this thesis is based on a much more limited set of questions and we rely on the assumption that correct answers to such questions are likely to be correlated with knowledge of other issues too. Nevertheless, it is worth noting that some voters tend to be "specialists": instead of being broadly informed about the main political issues, they learn only about issues that are of direct relevance to them. More generally, the greater relevance of some issues to specific sub-population groups makes such groups more likely to be aware of them. Delli Carpini and Keeter report that blacks were more informed than whites on racial issues and females more informed on gender-related issues (in spite of both blacks and females being less informed than average on other issues).

Another very important fact about the distribution of political knowledge is its clear relation with a number of observable individual characteristics like education, gender, and race. A number of possible explanations can be given to such a relation; these range from resource availability (to buy and process information) to historical exclusion of some groups from political life, which results in a lack of civic capacity[18]. To understand the causes of this unbalanced distribution represents the first step towards finding the way to promote more effective political equality. This is especially important if we think, as

[17] Page 174.

[18] For further evidence and discussion of this phenomenon see Delli Carpini & Keeter (1996; 135-177).

will be argued in this thesis, that the consequences of such unbalances are likely to be transferred into the area of public policy-making, where some groups could find more adequate representation than (and even at the expense of) others.

1.4 Information and voting behaviour

To be able to discuss the relationship between information and public policy we first need to understand what the impact of information on voting behaviour is. Recent literature in both economics and political science has contributed towards this aim. A number of theoretical studies have focussed on political participation, showing that having more detailed knowledge of candidates' characteristics and proposals could increase the chances of participating in political life by increasing the utility associated with electoral choices. In the context of elections, participation has mainly been identified with turnout. At the same time, it has also been argued that information might increase the "quality" of participation, and a second stream of studies has tried to understand whether better informed citizens are more responsive to electoral platforms, can be expected to take better decisions, and therefore to extract better outcomes out of the political process. It is clear that these two aspects of the influence of information on voting are not disjoint. Participation has to do with responsiveness, and it is an aspect of voters' response to candidates' proposals. We make such distinction mainly for exposition purposes, and because most literature seems to have dealt with the two quite separately.

1.4.1 Information and participation

"Government by the people" is the meaning of the Greek word "democracy": thus, the definition of democracy in itself implies people's participation in public decision-making. Although not many scholars would contest this definition, opinions about the type of participation that democracy requires are quite diverse. Two broad attitudes towards participation can be found in contemporary democratic theory[19]. On one side, champions of participatory democracy like Russeau or John Stuart Mill tend to associate the quality

[19] See Pateman (1970).

of democratic governance in a polity with the degree of participation of its members in the decision-making process[20]. In ancient Greek towns this was achieved via direct democracy. In modern systems, where direct democracy seems unfeasible, this would require a population that is well informed, participates in electoral meetings, stimulates and criticizes public officials, and turns out to vote in elections in large numbers. On the other side, proponents of a more limited notion of participation, like Schumpeter and Sartori, see citizens better placed as "controllers" of public officials than as directly taking part in decision-making[21]. Thus, high turnout rates are not necessary: the success of the system is only judged by its policy outputs.

One question however remains crucial, even for advocates of the second model: whether high participation might actually increase the quality of control imposed on elected representatives. This is an empirical question, and as such needs to be studied by linking positive models of elections and electoral behaviour with data analysis. Any attempt to answer this question cannot avoid considering the role of information in elections. If participation comes out of misunderstanding, propaganda or legal obligation then there is no need for it to improve the quality of governments. Indeed, both sides in the debate would agree that there is the need for at least "some" citizens to be informed, although they may disagree on the size of the "some".

A number of empirical studies has established a strong correlation between electoral turnout and a number of individual and systemic characteristics. Wolfinger and Rosenstone (1980) show that turnout is strongly predicted by a number of individual demographic variables. Education typically displays the highest influence, followed by income, age, marital status, and occupation. These results have been systematically confirmed by most subsequent studies, independently of the particular election examined. In addition, a number of studies have also shown that disposition variables such as party identification, sense of civic duty, and so on, affect the level of participation. Systemic characteristics also play a role: in particular, election closeness, registration laws, and local socioeconomic conditions[22] (average income, unemployment rate etc.).

Most empirical studies tend to show that costs and expected benefits of voting matter.

[20] See Russeau (1762) and Mill (1835 and 1861).

[21] See Schumpeter (1942) and Sartori (1987).

[22] See for example Patterson and Caldeira (1983), Cox and Munger (1989) and Leighley and Nagler (1992).

These could be regarded as good evidence in favour of the rational behaviour hypothesis. On the other hand, the negligible probability of being decisive that a rational and instrumental voter should recognize, should actually lead her not to vote at all (Downs, 1957; Riker and Ordeshook, 1968). As Aldrich (1993) argues, the costs and benefits of voting are low and this makes it possible that changes in expected election closeness, registration laws, weather conditions etc. have an effect on individual behaviour, in spite of not changing her expected utility by a great deal. If this is the case then comparative static analysis can be insightful in spite of the paradox of voting. Moreover, a large body of research clearly shows that, even in very simple situations, perceived probabilities are not necessarily those that correct estimations would deliver[23]. As will become clear, this is quite important for the methodology adopted in this thesis.

Is it possible to link stylized facts about turnout in a coherent theory grounded on rational use of information? In spite of the frequent allusions in the literature on voting behaviour about the possible role of information in driving turnout, there are still very few attempts to formalize this causal link. Ledyard (1984) presents a model of spatial electoral competition where voters are uncertain about the preferences and the cost of voting of other voters, and where abstention is admitted. Voters play a Bayesian game taking as given the two candidates' positions. Turnout would be positive for differentiated candidates' positions; electoral competition, however, leads candidates to convergence and this drives the equilibrium turnout to zero. Although Ledyard introduces turnout and asymmetric information in spatial models, his purpose was not really to explain how information affects turnout.

Feddersen and Pesendorfer (1996) consider a population where the level of information about relevant states of the world is exogenously determined. They consider two fixed alternatives (A and B) and two states of the world 1 and 2. Some voters always prefer A and some B independently of the state of the world (partisans), while others (independents) prefer A in state 1 and B in state 2. Some agents receive (costless and randomly) a message about the state of the world and this exogenously divides the population between informed and uninformed voters. At this point a decision on whether to implement A or B is taken by majority voting. In this model agents understand that their vote only matters if it is pivotal, therefore voting choice is conditional on the event of being decisive in the

[23] See for example Uhlaner and Grofman (1986).

election. Non-informed independent voters have an incentive to delegate their vote to the better informed to increase the chances of an informed aggregate decision: delegation is via abstention, which increases the probability of any informed independent voter to be pivotal. It is worth noting that a crucial assumption of this model is that all independent voters share the same objective function: thus, in delegating the decision to other independent voters no one incurs the risk of leaving the decision to people with different tastes. Moreover, partisans play a merely passive role: what matters in the model is only their (uncertain) number.

Feddersen and Pesendorfer (1999) generalize this model in several directions: first, agents are spread continuously across the political spectrum rather than sharing common values within groups; the candidate's quality index is not limited to two possibilities, so that partisanship can be (to a certain extent) traded off with quality; finally, the level of information held by voters can also vary and there is no presumption that a subset of voters is perfectly informed. However, some agents get noisy signals about the candidates' quality and this divides the population between those that have updated information and those who know only the ex ante probability of various quality realizations. This model finds again that more informed voters are more likely to vote than less informed ones. However, and quite surprisingly, increasing the fraction of informed voters in the population results in increased abstention.

Direct evidence[24] on the link between information and turnout is still very limited. Delli Carpini and Keeter (1996), using data from the 1988 NES Survey, show that political knowledge is a good predictor of electoral turnout, controlling for a number of individual characteristics. Sanders (2001) presents evidence on the 1996 US presidential election and shows the importance for turnout of perceived uncertainty about candidates. This effect is also linked to preference intensity.

Taking information as given in a model of turnout neglects the incentives that people face in acquiring information. In theoretical terms, if we think it is appropriate to use rational choice theory to explain voting behaviour, then it is legitimate to expect voters to apply rational calculus in the phase of information acquisition. This means that citizens are not just randomly informed and that, as an abundant empirical literature makes clear,

[24] I.e. evidence not based on variables that are just correlated with information, like education and the like.

political knowledge is correlated with a number of other individual characteristics that ultimately affect political preferences (e.g. income).

On empirical grounds such limits are not less serious: when estimating a turnout equation having indicators or proxies of political knowledge on the right hand side there is a serious possibility of capturing a spurious correlation[25]. A number of unobservable variables might affect both information acquisition and political participation and simple regression could deliver biased coefficients. Thus, a theory of information acquisition is also necessary in order to overcome this problem, as it can provide the appropriate instrumental variables for political knowledge. An analysis in this direction will be presented in chapter 2.

Endogenous information is introduced in Matsusaka (1995), that presents a decision-theoretical model of costly information acquisition and turnout. The utility from changing the election outcome is higher when the voter is more confident that she is voting for the candidate she actually prefers. Information, through Bayesian updating, increases such confidence and therefore also makes citizens more likely to vote. Matsusaka goes further by relating his results to a number of empirical regularities found in previous studies. Education and age, for example, reduce the cost of acquiring information and are therefore positively correlated with turnout; so does campaign spending or being contacted by a campaign worker before an election.

Disposition variables like party affiliation and other measures of ideological motivation are normally good predictors of turnout. There are, however, some good reasons to be cautious about their usage in empirical investigation. In answering questions about their preferences, attitudes and sense of efficacy it can be the case that respondents simply rationalize their behaviour, thus not providing any real insight on the link between such variables and behaviour in itself. It has been observed, for example, that responses can vary substantially with the order in which questions are posed (see for example Bishop, Oldendick, and Tuchfarber, 1984, and Abramson, Silver, and Anderson, 1987). Nevertheless, it is reasonable to expect that ideology plays an important role in determining turnout. Palfrey and Poole (1987) present an empirical study of the relationship be-

[25] Delli Carpini and Keeter (1996), for example, suggest that "the existence of a strong interdependence between knowledge and participation has an important practical implication: efforts made to boost one of these will, in all likelihood, benefit the other". This is a potentially wrong conclusion derived by interpreting correlation as causality.

tween information, ideology, and voting behaviour. Using ICPSR survey data from the 1980 presidential election in the US, their analysis shows that information is significantly related to both political extremism and turnout. Voter information is positively correlated with ideological extremism and negatively correlated with indifference between candidates. Moreover, not only are more informed citizens more likely to vote, but their vote is also more predictable, in the sense that they exhibit less randomness when voting behaviour is predicted using political preferences[26].

One element in this picture has so far been neglected, namely the role of ideology in determining information acquisition. There are a number of reasons, for example, for why people might tend to expose themselves to information sources they trust, thus with their same ideological bias. Berelson *et al.* (1954) and a number of subsequent studies found that voters practice selective exposure to information, for example simply paying more attention to information that is favorable to their own views. A theory of information acquisition should take into account that prior beliefs matter and that people do not share common priors on candidates. Thus, ideology should affect voters' knowledge of political matters and influence participation twice: directly, as it affects the perceived difference between candidates, but also indirectly, through its effect on the acquisition of political information. In chapter 2 we will introduce ideology in our model of information acquisition and provide evidence on its effects on both political knowledge and turnout.

1.4.2 Information and responsiveness

Are better informed voters more responsive to electoral platforms than less informed ones? The answer to this question is of extreme importance for our comprehension of the functioning of electoral systems. If the answer is yes then electoral competition should tend to deliver policies that are more favorable to the informed; this is especially relevant as political awareness is often correlated to a number of variables that also tend to be related to policy preferences.

Another possibility is that uninformed voters manage to behave as if they were informed, thus again making information not too relevant in terms of electoral outcome. A

[26] Preferences are recovered either by respondents' self-placement on a liberal-conservative scale, relative to their placement of candidates, or from self-placement on a number of issues like defense spending, inflation or government aid to minority groups.

vast literature, using different arguments and models, argues that citizens use heuristics that are both sufficient to make reasoned choices and cost effective. Sniderman, Brody and Tetlock (1991) define heuristics as "judgmental shortcuts, efficient ways to organize and simplify political choices, efficient in the double sense of requiring relatively little information to execute, yet yielding dependable answers even to complex problems of choice".

Using spatial models of elections, a number of papers[27] show that it is possible for imperfectly informed voters to emulate the behaviour of the better informed ones when they receive a credible signal from a source whose preferences are known. Sobel (1985) shows how informative equilibria may arise from repeated interactions when the information provider is initially not perfectly credible. It is then clear that if party platforms tend to be consistently tied to the interests of specific socioeconomic groups, then party identification can provide a simple and effective way to vote instrumentally.

In the context of heuristic decision-making, parties, pressure groups, opinion leaders etc. have a crucial role in transmitting simple and effective information to voters. Popkin (1991) argues that candidates' positions on most issues are correlated between them or to other variables. Therefore, it is only necessary to be aware of candidates' stands on few variables to make accurate inferences on the whole spectrum of issues and correctly estimate expected utility of voting for each candidate. Wittman (1995) argues that voters' costs of becoming informed have been vastly exaggerated; in particular, the returns to political entrepreneurs and lobbyists give them enough incentives to make information available to voters, leaving these with little or no cost to bear. Moreover, following an analogy with markets, Wittman stresses that competition among these *advisers,* ensures that voters get enough information to take the right decision.

Lupia and McCubbins (1998) focus on the role of institutions as "informational crutches": "it is not enough, for the success of democratic delegation, that institutions alter the incentives of democratic agents. Rather, it is also necessary that the incentive-altering effects of institutions make agents and speakers trustworthy *and* that democratic principals perceive institutions to have this effect. (...) Democratic institutions can, in this way, establish the conditions for persuasion, enlightenment, and reasoned choice". If

[27]See Calvert (1985), McKelvey and Ordeshook (1984), Grofman and Norrander (1990) and Lupia (1992).

clarity of interests, threat of verification of signals, penalties for lying and costly effort are all elements that the literature on signalling and cheap talking has identified as important for signals to be informative, then institutions that increase them will also increase the amount of information that citizens can extract from simple cues like party affiliation, endorsements, campaign spending.

Evidence on the use of shortcuts by voters has been provided by Brady and Sniderman (1985): using the US National Election Study, they show that voters tend to infer the relationship between what candidates prefer and their own preferences from information provided by particular trusted groups. McKelvey and Ordeshook (1984) provide experimental evidence on the use of polls about the opinions of groups whose interests are known. Lupia (1994) analyzes a survey of five complex insurance reform initiatives in California to conclude that "access to a particular class of widely available information shortcuts allowed badly informed voters to emulate the behaviour of relatively well informed voters".

Other scholars, however, are less optimistic about the capability of badly informed voters to act instrumentally. Mondak (1994) points to experimental evidence on the higher likelihood of mistakes by agents that use simple heuristics. More direct evidence against the "equivalent behaviour" hypothesis is provided by Bartels (1996): using the National Election Study surveys conducted for six US presidential elections, he finds that uninformed voters "do significantly better than they would by chance, but significantly less than they would with complete information, despite the availability of cues and shortcuts". In general, it is still possible to argue that more information is better even in the context of simple heuristic decision making . This point is made clear by Delli Carpini and Keeter (1996): "The distinction between the 'civic knowledge' and the 'heuristic' schools is less sharp than often supposed. (...) The use of shortcuts describes a human condition rather than a particular form of decision-making. The issue then, for both schools of thought, is not whether we use partial information to make political decisions, but the reliability and validity of the specific information we do use".

In short we can say that the heuristic decision-making argument shows that observed levels of political knowledge are not necessarily alarmingly low, as many scholars tend to assume: people acquire only the information they need and, under appropriate circumstances, this is enough to allow "ordinary citizens" to make "extraordinary choices".

Nevertheless, this does not imply that having more information gives no advantages in the political arena or allows better control over delegates. In particular, the argument that democracy works well with little information seems better applied in cases where accountability is at stake: but if different principals have different preferences then the issue is not just one of control over delegates but also one of redistribution. Other things equal (i.e. verifiability conditions, penalty for lies etc.) it is likely that more information can be extracted from the same environment by certain groups rather than others. Given that knowledge of others' preferences, interests, and competence is necessarily imperfect, political knowledge and its distribution is very likely to affect final electoral outcomes.

1.5 Consequences for public policy

Concluding this survey, we come back to the original question: does political information affect public policy-making through its impact in elections? As we have tried to clarify so far, the answer to this question depends on the answers we give to a number of related questions. Are voters capable of understanding electoral platforms and selecting the candidates that would better serve their interests? Are they capable of monitoring elected representatives? What do they know about the link between public policies and final outcomes? Do candidates try to learn about their constituents' preferences? Do they translate voters' preferences into implemented policies? This involves, among other things, the functioning of election mechanisms, the effects of information on voting behaviour, and the relationship between electoral politics and public policy.

One first obvious implication of what has been said so far is that if information does not affect voting behaviour then it should have no impact on public policy. However, it is also possible to argue that information has no effect on policies in spite of having an effect on voting behaviour. Formal models of elections have identified ways for the lack of detailed information by some (and even many) voters to have limited impact on aggregate choice. The simplest possible way is to assume that uninformed voters make mistakes, but that mistakes are just white noise in the election process. In other terms, if uninformed voters have equal probability to make mistakes in any direction, then this should not affect majority voting outcomes. This is a simple way to obtain "full information equivalence", i.e. the possibility for a majoritarian system to aggregate individual pieces of information

into the same final outcome that a perfectly informed population would choose. As we have already discussed, this idea can be dated back to Aristotle; Condorcet (1785) and recent formal models of collective decision-making have clarified the conditions required for this to happen[28]. Problems with this argument, however, arise if errors are correlated: if, for example, there is systematic bias in press coverage, then it seems more likely for voters' mistakes to be concentrated on some specific options. In this case, differently from Condorcet, the population size does not help in canceling out errors[29].

Feddersen and Pesendorfer (1996, 1997 and 1999) have been extremely influential in supporting full information equivalence. Conditioning on the event of being decisive reveals a lot of information to uninformed voters. Hence, although behaving differently from the informed[30], the uninformed manage to get their preferred policy by being adequately sophisticated. The relevance of these conclusions to actual elections have been seriously questioned. Margolis (2001 and 2002), for example, is very skeptical on this point: "results are contingent on a combination of very specific common knowledge interacting with an effectively leak-proof arrangement of socially relevant but strictly private information"[31].

As noted previously, it should not be overlooked that, even if people use the little information they have in the best possible way, this does not imply that, in comparative terms, more information would not be better. In the case of Feddersen and Pesendorfer, their logic can be pushed to cases where only one informed voter would be enough, in a symmetric equilibrium, to deliver the right choice. However, for information aggregation to occur, at least some information must be available: to be able to act strategically, the "uninformed" voters need accurate prior knowledge about other agents' preferences and distribution, which is more than one could expect from voters in an election. On top of this, at least some voters (the informed) should receive informative signals. Thus,

[28] Condorcet (1785) considers a population of imperfectly informed agents with a common objective function having to choose the best of two outcomes by majority rule. He shows that the probability of a correct choice increases with the population size. See also Young (1988) and Austen-Smith and Banks (1996). Feddersen and Pesendorfer (1998) show how the possibility of having full information equivalence depends on the decision rule: unanimity, for example, decreases the chance of having an informed collective choice.

[29] Berg (1993).

[30] As seen in the previous section, for example, uninformed voters are more likely to strategically abstain to increase the probability of an aggregate informed decisions. Also, the uninformed are more likely to vote against their signal if they cannot abstain.

[31] Margolis (2002).

even in a world of extremely sophisticated agents, electoral systems can work with little information, but not without information at all. Finally, in this literature the reliability of signals is not questioned. In real elections voters have virtually never the possibility to get first hand information as everything is filtered by opinion leaders, organizations, and, above all, the media; in real elections we definitely lack an objective and universally trusted signal. When signalling can be strategic, cheap talk games clearly show how a commonality of interests with the signaller (and therefore some knowledge of her preferences) is a condition to get informative equilibria.

The impact of citizens' political knowledge on public policy depends not only on how it affects their voting behaviour. Asymmetric information changes the incentives faced by politicians in designing electoral platforms and in delivering the promised policies. It can be useful on this point to distinguish between two dimensions that concern the link between elections and policies.

The first, which can be called "vertical delegation", concerns the capability of electoral processes to select the best candidates and then make elected representatives accountable to other citizens. The centre of attention is on the conflict between the interest of decision-makers and that of the polity. Elections, in this case, serve as screening and disciplining mechanisms; the perspective of future elections should provide incentives for decision-makers not to abuse of their power for private purposes.

The second dimension is one of "horizontal conflict". Citizens have different preferences for public policies and, unlike for goods allocated by markets, voluntary exchange and the price mechanism cannot help in this case. Therefore, if and when a public decision is reached, it is binding for everyone. The conflict among different preferences must be solved by centralized decisions that inevitably involve some form of (implicit or explicit) redistribution.

A good framework for examining accountability of public officials is the agency model. Since Barro (1973) and Ferejohn (1986), agency models have been used to study the incentives faced by an incumbent (agent) to provide the policies preferred by citizens (principals). This is natural since an agency relation is concerned with some idea of performance. Examples of applications of agency models to politics include policy manipulation for re-election purposes (Harrington, 1993), political business cycle (Rogoff, 1990), yardstick competition in tax setting (Besley and Case, 1995a), and the form of

transfers to special interest groups (Coate and Morris, 1995).

The main conclusion in this class of models is that the perspective of being re-elected reduces the rent extracted by incumbents in equilibrium. This is what Banks and Sundaram (1999) call the "performance effect". There is also a "selection effect", in the sense that not all agents will be equally likely to be re-elected: good types (more competent or with motivations more similar to those of citizens) will have generally a better chance to be confirmed in office. Thus, the perspective for the incumbent of facing re-election gives a double advantage to citizens: it delivers both better policy-making by the incumbent in office (independently of his type) and a higher probability to have a better type of policy-maker in service in the following period. Besley and Case (1995b) use data on gubernatorial administrations in the US to provide evidence in support of the first claim[32]. Hence, one first important conclusion is that elections are, at least to a certain extent, a good device to affect public policy in the direction desired by the majority of citizens. The perspective of facing a re-election works as a commitment device.

Information plays a key role in all agency relations: thus, when decision-making power is delegated to governments, information availability is crucial for accountability and therefore for public policy. Voters can be badly informed about the incumbent's competence, her implemented policy, or the link between policy and outcome. In each of these cases voters will implement a different re-election strategy, either based on past policy if this is observable (Rogoff and Sibert, 1988; Rogoff, 1990), or on past outcome (Ferejohn, 1986; Alesina and Cukierman, 1990). Harrington (1993) takes up the issue of voters' uncertainty about the efficacy of different economic policies. He finds that the more uncertain voters are about the efficacy of policies the bigger the impact of performance on the electoral outcome and the larger the possibility of policy-manipulation open to incumbents.

One special characteristic of the political market is that almost all the information available does not come from direct observation or knowledge of facts, but rather from what the few informed people say. Claims are sometime non-verifiable and many other times the verification is too costly for any single citizen to pursue, considering also the well known collective action problems of large elections. If, however, for any reason (e.g.

[32] Governors face a term limit of two mandates. Therefore, in each given moment there are governors facing the perspective of elections and "lame ducks" that will terminate their service anyway. Besley and Case find evidence of different behaviors in the two cases.

instrumental voting, expressive benefit, pure entertainment, or else) there is a demand for such information then special firms, the mass media, should emerge to "produce" and sell it to interested customers. It is indeed not surprising to observe that the historical development of the newspaper industry has proceeded hand in hand with the development of democratic institutions.

Recent theoretical and empirical research, mostly still unpublished, has tried to disentangle the role of information supply in making public officials accountable to citizens. In Besley and Prat (2001) the media (a number of competing outlets) can derive their revenue from two sources: audience (and therefore advertising) and a transfer from government (in the form, for example, of favourable regulation) in exchange for silence on possible bad news about the incumbent. The higher the number of outlets accepting the deal with the government, the larger the potential audience from breaking news for the non-captured outlets. Therefore, to keep every outlet silent, the government must pay each of them as if it was the only one that could break news. Hence, a rise in the number of outlets makes it more expensive for the government to capture the media. Since in equilibrium the government can only either pay everyone or nobody (just one outlet breaking news is enough for the citizenship to be informed), increasing the number of outlets renders less likely the possibility of media capture. Media pluralism is, therefore, good for information availability. Besley and Prat also provide cross-country evidence by linking foreign media ownership with corruption indices. A foreign owner can be less prone to be captured because, for example, has less economic interests under government's regulation: this increases the transaction cost of transfers. Djankov *et al.* (2001) find that the number of state owned newspapers is also a good predictor of corruption. Besley and Burgess (2002) provide evidence on Indian states responsiveness to calamities and find that this is associated with the circulation of newspapers and electoral turnout. This idea dates back to Sen (1981, 1984) who pointed out the role that newspapers can play in preventing famines, by increasing citizens' awareness and therefore government activity in prevention.

All this seems to suggest quite clearly that information and mass media matter for policy outcomes. It is interesting to note that even advocates of "limited information politics", like Lupia and McCubbins, agree on the role of the media, and include them among the "informational crutches" that allow ordinarily (i.e. little) informed citizens

to make the right choices: "the existence of a free press increases the likelihood that campaign statements will be verified. While competition can create multiple potential verifiers, freedom of the press provides these verifiers with an avenue whereby they can provide public verification of political statements. (...) These institutional features can increase the threat of verification and can thus increase the likelihood that voters will be capable of reasoned choice"[33].

Coming to the horizontal dimension, the link between information and policy is driven by voters' responsiveness to electoral platforms, which implies both to be able to distinguish the most favorable proposal and to turn out to vote for it. "The very groups who are disadvantaged economically and socially are also less politically informed and, thus, disadvantaged in the struggle over the political allocation of scarce goods, services, and values. (...) The more informed one is, the more likely one is to send clear, policy-oriented messages to political elites (...). To the extent that political elites respond to such signals - a central tenet of any theory of representative democracy - informed citizens are likely to have their concerns taken more seriously"[34]. Empirical evidence tends to show that turnout is often the transmission chain of this mechanism. First, voters and non-voters systematically differ in their socioeconomic and demographic characteristics and therefore in their needs and policy preferences[35]. Second, most evidence suggests that "low voter turnout means unequal and socioeconomically biased turnout"[36]. Thus, if we expect voters' preferences to be represented in policies, turnout levels should determine policy outcomes.

Empirical evidence that aggregate turnout is a predictor of welfare spending has been provided by Peterson and Rom (1989) for US states and Hicks and Swank (1992) for industrialized countries. Lindert (1996), analysing a panel of OECD countries, finds that "a stronger voter turnout seems to have raised spending on every kind of social program, as one would expect if one assumed that the social programs cater to the lower income groups whose voter turnout differs most over time and across countries". Hill and Leighley (1992) and Hill, Leighley, and Hinton-Andersson (1995) use US survey data to derive aggregate measures of turnout by social class and combine them with state-

[33] Lupia and McCubbins (1998).
[34] Delli Carpini and Keeter (1996).
[35] Wolfinger and Rosenstone (1980), Verba *et al.* (1993).
[36] See Lijphart (1997).

level data to provide direct evidence of the effect of lower-class mobilization on welfare spending. Using US state-level data for the years 1950-1988, Husted and Kenny (1997) show how the extension of the voting franchise (thus favouring participation by the poor and the minorities) has caused an increase in welfare spending, leaving all other spending unaffected. If turnout can be explained also by information, as we have seen in the previous section and as Chapter 2 will make clearer, then we are facing more than a presumption of the impact that information could have of public policy. Chapter 3 of this thesis will present an information-based model of electoral competition and voters' responsiveness that provides a possible explanation for these findings.

Another source of differentiated influence on the electoral process can be the mass media. This mechanism is studied by Stromberg (2002): mass media derive their revenue from advertising and some people are more valuable than others to advertisers (those who tend to consume more, like the richer, better educated, younger etc.). Hence, these people will be targeted by the media and office-seeking politicians will also design policies more favorable to media users, as those are more likely to be informed on platform proposals. Equilibrium policies can therefore be substantially altered by the functioning of the media market, independently of any ideological bias that the media could possibly throw into the electoral battlefield. Stromberg (2001) also provides evidence of this effect from the New Deal relief programme implemented in a period of rapid expansion in the use of the radio. He finds, controlling for variables that account for the needs of different counties, that the radio had a large and significant impact on funds allocation. Chapter 4 of this thesis will offer some theoretical foundations and evidence of instrumental behaviour by voters in acquiring information and, in doing so, will also provide evidence of newspapers' discrimination in information supply.

In conclusion, new research is increasingly providing both theoretical rationales and empirical evidence on the effects of the information market on public policy-making: it seems then clear that if we think it is worth opening the black box of policy formation then information cannot be left out of the picture. Although much is still to be done to gauge the impact of information on policies, it seems worth exploring the incentives faced by agents in the information market and their interactions with electoral institutions and the incentives provided by democratic political competition. This should lead to a better understanding of public policy and, in perspective, to new normative implications.

1.6 Thesis overview

This chapter has discussed the main motivations that stimulated this thesis. It also aimed at presenting the intellectual framework of this research, underlying some of the many unclear facets in our understanding of the determinants and the consequences of voters' political knowledge. The next three chapters aim at contributing to this research agenda by mainly focussing on the information demand side.

Chapter 2 studies the impact of information on electoral turnout. First of all, it introduces a model of political information acquisition that will be used, in different ways, across the whole thesis. Information acquisition is modelled as an individual production activity. Voters are endowed with a personal information production technology whose inputs are mass media and time. Exogenous individual characteristics and news availability determine the productivity of different agents, i.e. their capability to acquire, process and retain relevant political information: optimal individual media usage depends on such exogenous parameters, consistently with a number of stylized facts. We also introduce ideology, in the form of prior beliefs about candidates, and assume agents are heterogeneous in such beliefs. Our theoretical model predicts that information increases the likelihood of turnout for voters who are sufficiently independent, in the sense of not having a strong prior preference for one of the candidates. Information demand turns out to be a non-monotonic function of ideology: the voters with the lowest incentives to acquire information are those with the weakest (indifferent) and the strongest (partisan) prior beliefs. Our model predicts instead that the best informed citizens are those who moderately support one of the candidates. It is also shown that people with strong prior beliefs should reduce their likelihood of voting when informed. Thus, having extreme ideological priors affects turnout twice: first, directly, by increasing the preference for one of the candidates; second, via information, by reducing the probability of learning about candidates.

These theoretical results are tested using survey data on the 1997 general election in the U.K. The information acquisition model performs very well when confronted with data. A number of variables affecting the benefits and costs of acquiring information are good predictors of political knowledge: some of those have been traditionally used in empirical political science research, others have been added thanks to our theoretical

investigation. Education and news availability on the media emerge as good predictors of knowledge. Information is then shown to be a very good explanatory variable for turnout. By using appropriate instrumental variables, it is possible to show that this is not just a spurious correlation driven by unobserved heterogeneity: political information affects the probability of turnout in a very precise sense. Predictions on the impact of ideology on both information acquisition and turnout also receive good support from data, especially when ideology is measured by party identification.

This chapter leads us to conclude that information matters for turnout. It also matters for good collective decision-making. The model presented has only a "vertical" dimension of candidates: these can therefore be unanimously ranked. Information increases the chance of an enlightened choice by increasing the likelihood of turnout of independent informed voters. Also, agents with strong priors are those more likely to vote in the "wrong" way: thus, information increases the chances of an aggregate informed choice by increasing the probability of abstention of the extremists. Although derived within a different modelling frame, it is clear that our results may have some relevance for the literature on information aggregation.

Our empirical investigation shows quite neatly that political knowledge is increasing in income. Chapter 3 studies the implications of this fact for redistributive public policy. The focus is on individual incentives to acquire information. We argue that, even when the probability to cast a decisive vote is virtually negligible, voters have still private incentives to be informed about politics. Monitoring of current policy and expectations over future policies can be very useful for a number of private decisions like financial investments, choosing between public and private education, or the choice of a pension scheme. Thus, political information has also purely private returns. Under quite mild assumptions it is possible to show that information demand for private purposes is increasing in income. Office-seeking politicians should take this into account when competing for office. We study these effects within a unidimensional model of Downsian political competition. Our analysis carries three main implications for redistribution: 1) equilibrium policy does not converge to the median voter in the population but to the median informed voter; this means that redistribution is always less than predicted by traditional (i.e. with perfect information) Downsian models; 2) greater inequality in gross income distribution is neither necessary nor sufficient for more redistribution; inequality increases the redistributive

desire of the median voter but also the political weight of the rich, leaving the net effect undetermined; 3) restrictions on the policy space can have perverse consequences as they change the incentives of different segments of the population to be aware of public policy-making; for example, if the poor are not taxed and do not receive benefits they have less instruments and less incentives to monitor public policy-making.

Chapter 4 studies the political information market, by explicitly introducing mass media. It takes up two issues: voters' instrumental behaviour and the mass media bias on public policy. If voters act instrumentally on the political market, then they should be responsive to a number of incentives that can be picked up by observable variables. In particular, a close election increases the probability for each voter to be pivotal and should therefore increase the incentives to be informed. At the same time, the mass media can increase their profits by discriminating between areas with lower and higher political information demand: since price discrimination is not normally practised, they will discriminate in the supply of news. We will provide evidence on both voters' and mass media behaviour by focussing on electoral constituencies in Britain during the electoral campaign of 1997. We find evidence of voters' higher demand for information in marginal constituencies. At the same time, the mass media tend to supply more news to marginal constituencies but also take into account a number of other local characteristics that make readers more or less valuable to advertisers. Thus, they introduce a bias in information supply that is completely independent of potential ideological biases. This opens the possibility that policy-makers, following the information flows, will themselves discriminate between different constituencies and different population segments.

Chapter 2

Information Acquisition, Ideology, and Turnout: Theory and Evidence from Britain

2.1 Introduction

In an idealized vision of democracy, public decisions are the expression of preferences and opinions of all the members of a polity. In reality, modern democracies simply delegate decision-making power to one or some representatives. Elections represent the cornerstone of delegation and control. Thus, a modern version of participatory democracy would require a population that is well informed, participates in electoral meetings, stimulates and criticizes public officials, and turns out to vote in large numbers. This consideration leads many people, from political scientists to journalists, professional politicians, and simple citizens, to attach a special value to electoral turnout as reflecting how representative public decisions are of citizens' interests. Low electoral turnout is often seen as a symptom of little attachment to public matters and even to democracy in itself and is therefore associated with the danger of a self-referential political class.

Nevertheless, high participation rates are not necessarily symptoms of healthy democratic institutions, and it is possible to dismiss the danger of low turnout on a number of grounds. Abstention could simply be a sign of consensus to others' choices; abstainers in this case are not passive and disinterested, and low turnout would not signal anything else

than happiness with the *status quo*. Proponents of a more limited notion of participation, like Schumpeter and, more recently, Sartori, see citizens better placed as "controllers" of public officials than as directly taking part in decision-making[1]. In this case high turnout rates are not necessary: the success of the system is only judged by its policy outputs. Endorsing either vision of participation can lead to substantially different conclusions about the quality of our democracies and the possible steps to take.

Understanding the determinants of turnout is central to this debate and would give us substantial insights on the formation of public policy through the political process. An important step in this direction is to correctly evaluate individual motivations; in this sense, a long tradition of empirical studies that has mainly focused on aggregate data can give us only a partial picture. More recently, a number of empirical studies has established a strong correlation between electoral turnout and a number of individual and systemic characteristics[2].

Independently of which model of democracy one endorses, a crucial question remains whether high participation rates might actually increase the quality of control posed on elected representatives. Any attempt to answer this question cannot escape considering the role that political information plays in elections. Recent theoretical and empirical research seem to show that the circulation of political information increases elected public officials' responsiveness and accountability[3]. In general, as information plays a key role in agency relations, it is reasonable to expect information availability to be important for accountability (and therefore for public policy) when decision-making power is delegated to governments. On a different ground, the distribution of political information may have an impact on redistributive policy as office-seeking politicians will target their platforms at voters that are more likely to be aware of them[4].

Either implicitly or explicitly this literature assumes that information matters for voting behaviour. Nevertheless, as we have seen in Chapter 1, some authors argue that substantially uninformed voters manage to behave as if they were informed by using heuristics that are both sufficient to make reasoned choice and cost effective[5]. In brief, these theo-

[1] See Schumpeter (1942) and Sartori (1987).
[2] Wolfinger and Rosenstone (1980) has been the seminal work in this direction.
[3] See for example Besley and Burgess (2002) and Besley and Prat (2002).
[4] See Stromberg (2001 and 2002) and Chapter 3 in this thesis.
[5] See in particular Popkin (1991), Wittman (1995) and Lupia and McCubbins (1998).

ries tend to argue that knowledge of political matters is "behaviour irrelevant". Another stream of theoretical research argues instead in favour of "full information equivalence": a polity of substantially uninformed citizens can manage to reach the same outcome that a perfectly informed population would choose[6]. In this case uninformed voters behave differently from the informed as they tend to abstain in order to increase the probability of an informed collective choice.

The current debate suffers of at least two important limitations. First of all there is still no evidence of a causal link between information and turnout: some empirical research shows that there is correlation between the two, yet this is still far from implying that information determines participation[7]. Second, one important question has yet to receive the attention it deserves: what determines the extent of citizens' political knowledge? If we approach voting behaviour using the methodology of rational choice theory then there is no reason to leave political information acquisition out of our investigation. It will be immediately clear that the two questions are related.

This chapter will therefore try to address both questions, starting with the second. We will therefore introduce a model where the demand for political information is the outcome of a rational process, with its costs and benefits. This does not imply that we want or can "explain", in a strict sense, information acquisition: this would be a very difficult task, as it is hard to explain voting itself in large elections, given the negligible probability of each voter being decisive. However, as for most economic theory, our purpose is not much to explain why people desire something or why certain commodities are preferred to others, but rather how their demand and supply vary in accordance with relevant observables like prices, costs and institutional arrangements. Although we will try to spell out the basic motivations driving information acquisition, this agnostic approach constitutes the starting point of this work.

The first purpose of this chapter will be therefore to model information acquisition from a decision-theoretical perspective and to propose a theory of electoral turnout based on endogenous political information. Information acquisition will be modelled as an individual production *à la* Becker, where inputs are represented by mass media and time devoted to their usage. Different agents are endowed with different "technologies" to

[6] Feddersen and Pesendorfer (1996 and 1999).
[7] See Delli Carpini and Keeter (1996).

acquire and process information and are therefore able to grasp more or less information from the same exposure to media: in this sense we should expect a number of observable individual characteristics to be positively related to the capability to be informed. We will then proceed linking information with turnout and testing our results on British data.

Modelling information acquisition turns out to be quite important for the strategy of the empirical investigation. To see why, suppose we want to estimate the effect of information on turnout and let us consider the following simple model. A theory of instrumental voting starts from the idea that people vote in elections because they are interested in policies; in the classical formulation of Riker and Ordeshook (1968), a citizen votes if

$$PB + D > C \tag{2.1}$$

where P is the probability to cast a decisive vote, B is the gain in policy benefit derived from the victory of the preferred candidate as compared with the opponent, D is a psychic benefit to voting and C is its cost. Unfortunately, as it stands, this expression represents just a small progress from the tautological proposition that people vote if they like to do so. This consideration extends to information acquisition.

Political information helps the decision-maker to form a more precise idea about B, the difference in utility between, for example, two candidates. Other elements in the (2.1) can be influenced by information: the perception of P can for example be affected by published polls during the electoral campaign. We will focus on B, as the element that reflects the policy platforms, and can therefore be related to ideology and political information in a stricter sense.

We now assume that B depends on some decision to be taken by the elected candidate and that, in turn, this depends on a parameter δ; abbreviating we can write $B = B(\delta)$. The optimal decision for our voter should then take into account the value of δ. We can indicate with $W(\delta)$ the utility derived from the maximization of the equation $PB+D-C$: this is an indirect utility function. Now suppose that δ is unknown and has to be estimated for decision-making purposes. If $\widehat{\delta}$ is a more precise estimator than $\widetilde{\delta}$ of the true δ, then we can say that the value of using $\widehat{\delta}$ instead of $\widetilde{\delta}$ is given by

$$E_\delta W(\widehat{\delta};\delta) - E_\delta W(\widetilde{\delta};\delta) \tag{2.2}$$

where E_δ represents the expected value operator. If the cost of passing from the estimator $\widetilde{\delta}$ to the estimator $\widehat{\delta}$ is c (for example to acquire a larger sample of observations), then such acquisition will take place if

$$E_\delta W(\widehat{\delta};\delta) - E_\delta W(\widetilde{\delta};\delta) > c \tag{2.3}$$

However, political information can be acquired also for non-instrumental reasons. As in the voting equation, we can add a personal benefit b which represents some psychic enjoyment of political information, orthogonal to political preferences and observable relevant variables. The (2.3) becomes

$$E_\delta W(\widehat{\delta};\delta) - E_\delta W(\widetilde{\delta};\delta) + b > c \tag{2.4}$$

The problem in testing the effect of information on turnout arises as in practice D and b are likely to be correlated, both being driven by some sense of civic duty or pure enjoyment of politics. Finding a positive relation between information and turnout could therefore just be the consequence of omitting some relevant variables; this correlation does not reveal anything about the impact of information in itself on turnout.

A theory of information acquisition is therefore necessary if we want to identify the turnout model. This theory will provide the necessary instrumental variables to overcome this endogeneity problem and assess the impact of information *per se* on turnout. As a by-product of the empirical investigation we will also get a number of results about British citizens' political knowledge.

It should be noted that the value of information depends on the possibility (ex ante) that new observations induce a change in behaviour. This, in turn, implies that prior beliefs about the distribution of δ will affect our results. People have different beliefs about parties and candidates and we will call such beliefs "ideology". In general, apart from gathering information from mass media (or other sources), most people have their own prior opinions on political issues: these are reflected on both policy preferences and beliefs about how to reach given targets. These priors can be shaped by the influence of other people (e.g. parents), by personal knowledge and competence in political matters, by personal experiences etc. Such opinions or, in other terms, such ideological motiva-

tions, however formed, play an important role in voting decisions and on turnout itself[8]. Moreover, and more interestingly from our perspective, they can have an influence on the decision to acquire political information; thus, when we introduce endogenous information, ideology can influence voting both directly and indirectly via information. Another objective of this chapter is therefore to study such influences: the interaction between ideological motivations and information acquisition will lead us to new results on electoral turnout and will have implications for the way elections shape public policy. Differently from previous works, good collective choices may be favoured not only by abstention of uninformed voters but by abstention of informed voters as well.

This is admittedly a minimalist interpretation of the word ideology, that is instead often referred to broad theoretical constructions and general visions about politics and beyond it. For our purposes it will just be important that, in elections, these broad visions of the world will translate into different prior opinions about candidates. Although not pretending to have a theory of ideology, nevertheless we hope to provide new insights on the role of ideology in elections.

The chapter is organized as follows. The next section will introduce the main features of the theoretical model while section 2.3 will analyse the case of a non-polarized polity, when each decision-maker has prior beliefs that make her substantially indifferent (ex ante) about the candidates. In section 2.4 we turn to the role of ideology, thus considering the possibility of prior beliefs that attach different values to the various candidates. Section 2.5 gives some information about the 1997 general election in Britain, on which the theory is tested, and presents the datasets that are used in the empirical investigation. Section 2.6 outlines the estimation strategy, the links between the theory and the empirics and the way to solve potential endogeneity problems. Section 2.7 illustrates the results of the empirical analysis and Section 2.8 concludes the chapter. Further details about both the theory and the data can be found in the Appendices.

[8] Palfrey and Poole (1987) present an empirical study of the relationship between information, ideology, and voting behaviour.

2.2 The model

We consider a polity with two political parties I (incumbent) and O (opponent) and a set Ω of citizens who vote to elect a public decision-maker. The incumbent politician decides the value of a public policy parameter $a \in [0, \overline{a}]$. We assume a one-to-one relationship between politicians and policies: in other terms (abusing the notation) candidate a delivers policy a. The incumbent policy-maker a_I faces an opponent selected by party O. The opponent candidate selection process is unknown to citizens and will be represented by a probability distribution function $F_O(a)$,with corresponding density function $f_O(a)$. On the other side, when the politician in office implements her preferred policy a_I she reveals her type to citizens: therefore, while a_I is common knowledge, citizens do not know the opponent's type a_O.

Citizens' preferences over policies are represented by a utility function $V(a)$. We assume that all citizens have the same preferences over a and that $V'(a) > 0$. This is admittedly a strong assumption: people generally differ in their preferences over public policies. However, it is possible to regard a not as a specific policy dimension but rather as a more general measure of "good government". Indeed, it seems very often the case that citizens, rather than differing on their final aims, have different opinions on the most appropriate ways to reach those aims. Let us take the example of health care: not many politicians would claim they don't care about people's health. However, different strategies to reach good health services are rationalizable and are indeed rationalized during electoral campaigns. Another good example is gun control: both the supporters and the opposers of increasing gun regulation claim that their advocate policy would decrease criminality and increase the average citizen's safety. Both cases are logically possible and evidence is often not clear or easily manipulable. It turns out that in many cases heterogeneity is not as much a matter of preferences on final goods as it is in beliefs about the effectiveness of different policies. In this sense $V(a)$ can be taken as representing meta-preferences on good government. This is just a convenient simplification: heterogeneous preferences can be introduced at the cost of extra technical complications and little new insights compared with the present analysis.

Heterogeneity is instead introduced on prior beliefs. We assume that citizens have different prior distributions $F_O(a)$ about the opponent's type. This formulation is admit-

tedly non-orthodox, though, as discussed in the first chapter, heterogeneous prior beliefs are inevitably part of political life[9]. The set of admittable prior distribution functions is indicated with $\mathcal{F}$.

During the electoral campaign, citizens can gather information on candidates. In our case this is limited to the opponent, as the incumbent's type is common knowledge. Citizens are endowed with an information gathering technology that is representable by the probability $q(t, k|E, M)$ to learn the realization a. The inputs of this personal production function of information are an information source of quality $k \in \mathcal{K} \subset \Re_+$ (this includes newspapers, television, radio etc.), and time $t \in \mathcal{T} \subset \Re_+$ devoted to extract information from this source[10]. The opportunity cost of time t is represented by w, the marginal cost of quality of the information source is r.

This technology also depends on a vector of parameters E that affect the ability to extract and process information or the capability to use more sophisticated information sources. In empirical applications E will include variables such as education and age.

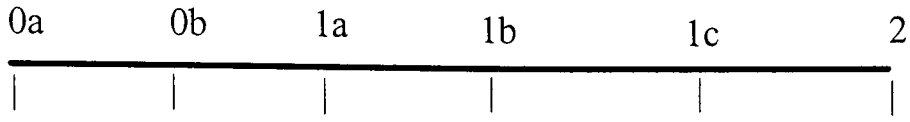
The probability to learn a_O depends on the concentration of news about a_O on the information sources with quality k; thus, q will also depend on a parameter M, that reflects information supply, typically via the mass media. There is however a difference between k (the source's quality) and M (information supply). The first can be individually chosen, according to each individual's interests and capabilities. M instead reflects the salience of given issues or constituencies on the media and is therefore independent of citizens' willingness to acquire information. We will assume that $q(t, k|\cdot)$ is always increasing in E and M.

At election time citizens compare the benefits of the two candidates: informed citizens will compare $V(a_I)$ with $V(a_O)$, the utility they derive from the realized opponent candidate; uninformed citizens will instead use their prior beliefs on the opponent's type. The benefit from voting is defined as the (expected) difference in utility from the two candidates, taking into account the probability that each voter has to be decisive. In the current analysis we will not consider any non-instrumental motivation for voting. This, however, is normally represented just as a constant, and therefore would not change our

[9] For a discussion of this assumption, see Harrington (1993).

[10] Note that k is just a quality index and does not represent in itself specific sources. We only assume that each specific newspaper, magazine, television channel or radio station can be mapped into the space $\mathcal{K}$.

Figure 2-1: Time Line



0 = Incumbent implements a_I and reveals her type

0a = Opponent selection from distribution F(a)

1a = Choice of t^* and k^*

1b = realization of $q \rightarrow \widehat{q}$

1c = election

2 = winning candidate implements her preferred policy:utility is realized

results.

Finally, voting is costly: we represent the cost of voting with $C \in \mathcal{C} \subset \Re_+$. We will consider a fixed C but nothing would change if instead we assumed that C was distributed across the population according to any given distribution function, as long as the distribution of C is independent of the distribution of prior beliefs about the opponent. Each agent knows his own C.

After the election, the elected politician implements her preferred policy a^*. The sequence of events is represented in Figure 2-1.

2.3 Information acquisition and voting

In this section we will first characterize the value of and demand for information. Then we will restrict our attention to the case of a non-polarized polity by introducing restrictions on prior beliefs and cost of voting.

We start by solving the model backward to characterize the information acquisition process.

At time 2 the winning candidate implements her preferred policy: that will be a_I if the incumbent is confirmed in office and a_O if the opponent candidate wins. For brevity we will indicate $V(a_I)$ with V_I and eliminate the subscript from the functions $F(\cdot)$ and $f(\cdot)$;

where there is no risk of confusion we will also use a for the opponent's type, eliminating the subscript. Abusing the notation we will also indicate with I the decision to cast a vote for the incumbent, with O a vote for the opponent and with A the decision to abstain. Focusing on turnout, we will also indicate with $T = 1$ the decision to vote and $T = 0$ the decision to abstain.

The decision problem of an uninformed citizen at the election stage is then

$$\max_{\{T\}} T(P|\int [V(a) - V_I]dF(a)| - C) = \widetilde{W} \tag{2.5}$$

where P is the (exogenous) probability to be a decisive voter.

For a citizen who knows the type of the incumbent the problem is instead

$$\max_{\{T\}} T(P|V(a) - V_I| - C) = W^*(a) \tag{2.6}$$

The ex ante value of an informed versus an uninformed decision is then given by

$$\Delta = \int [W^*(a) - \widetilde{W}]dF(a) \tag{2.7}$$

At the beginning of period 1 citizens decide about information acquisition. As mentioned, we assume they are endowed with an information gathering technology representable as the probability $q(t, k|E, M)$ to learn the realization a. We make the following assumption on $q(t, k|E, M)$.

Assumption 2.1 $q_t > 0$, $q_k > 0$, $q_{tE} > 0$, $q_{tM} > 0, q_{kE} > 0, q_{kM} > 0, q_{tt} \leq 0$, $q_{tt} \times q_{kk} - (q_{tk})^2 \geq 0$

This is just a simple assumption on the relationships between inputs and output. Indeed, we treat $q(\cdot)$ not differently from any standard production function.

Now we can state period 1 optimization problem for a generic citizen as:

$$\begin{aligned} & \max_{t,k} q(t, k|E, M)\Delta - wt - rk \\ s.t.\ t &\in \mathcal{T} \\ k &\in \mathcal{K} \end{aligned} \tag{2.8}$$

In Lemma 1 we prove that the expected value of information is positive.

Lemma 2.1 $\Delta \geq 0$

Proof: See Appendix 2A.

It is then straightforward to prove the following:

Proposition 2.1 *The optimal functions* $t^*(E, M, w, r)$ *and* $k^*(E, M, w, r)$ *are both increasing in* E, M *and decreasing in* w, r.

The demand for information, expressed as demand for mass media and time devoted to their usage, is then increasing in the technology parameter E and in the supply of relevant information by the media M. The same will be true for the probability Q to know a :

$$q(t^*, k^*|E, M) = Q(E, M, w, r).$$

In the rest of this section we will analyse the case in which prior beliefs are such that agents would not vote for any of the candidates if uninformed. In other words, ideologies are weak in terms of their effects in elections; we will also say that the polity is not polarized.

Assumption 2.2 $P|\int[V(a) - V_I]dF(a)| \leq C, \forall F \in \mathcal{F}, \forall C \in \mathcal{C}$.

Uninformed agents will therefore always abstain.

We can now prove some results on electoral turnout, starting by linking the probability to be informed to the ex ante probability to vote, i.e. the probability of voting before the actual type of the opponent is revealed. This ex ante perspective is indeed the only allowed for an external observer, at least if we want to maintain an agnostic view about the actual quality of candidates and their political distance.

In the following we will always assume that $P[V(\bar{a}) - V_I] > C$ and $P[V_I - V(0)] > C$.

Proposition 2.2 $\frac{\partial \Pr(T=1|Q)}{\partial Q} \geq 0$.

Proof. See Appendix 2A.

The probability of voting for any candidate is increasing in information, i.e. on the probability to know the opponent's type. It is also clear that our theory links the probability of voting to a number of individual and environmental characteristics. This provides a theoretical foundation for a number of well established stylized facts on turnout.

Proposition 2.3

$$\frac{\partial \Pr(T=1|E,M,w,r)}{\partial E} \geq 0, \frac{\partial \Pr(T=1|E,M,w,r)}{\partial M} \geq 0, \frac{\partial \Pr(T=1|E,M,w,r)}{\partial w} \leq 0, \frac{\partial \Pr(T=1|E,M,w,r)}{\partial p} \leq 0.$$

Proof. See Appendix 2A.

The capability to acquire information (as measured for example by income or education), as well as the amount of information supplied, both increase the probability that a citizen, *ceteris paribus*, will vote. Thus our theory can explain some of the common findings of most empirical research, like the positive correlation between education and turnout. At the same time we have new testable results linking the probability of turnout with the cost of acquiring information.

2.4 Ideology

In the following we will introduce ideological prior beliefs. The purpose of this section is to analyse how information demand and turnout depend on ideology. This will be accomplished by performing comparative static analysis under different hypotheses on citizens' prior beliefs. Although the word "ideology" has a much broader meaning and can be subject to various interpretations, it is natural to think of ideology in our model as deriving from prior beliefs about the opponent candidate. It is clear that, in the real world, ideology concerns beliefs about all candidates: however what really matters for voting decisions is the perceived position of one candidate relative to the other and to the cost of voting.

We will start by defining ideology according to citizens' beliefs.

Definition 2.1 (Weak Ideology) *A weak I-ideology (O-ideology) consists of prior beliefs F(a) s.t.*

$$\int [V(a) - V_I] dF(a) < 0 \quad (\geq 0)$$

It should be noted that, according to this definition, a citizen is weakly ideological only in relation to beliefs. This, however, does not guarantee that an ideological citizen will vote if uninformed: a more stringent definition of ideology would require prior beliefs to be such that the *ex ante* distance between candidates is sufficient to overcome the cost of voting. We introduce therefore the following definition:

Definition 2.2 (Strong Ideology) *A strong I-ideology (O-ideology) consists, for given P and C, of prior beliefs $F(a)$ s.t.*

$$\begin{aligned} P\int[V_I - V(a)]dF(a) &\geq C \\ (P\int[V(a) - V_I]dF(a) &\geq C) \end{aligned}$$

It is then possible to divide the set of prior beliefs $\mathcal{F}$ into three groups:

$$\begin{aligned} \mathcal{F}_I &= \{F(a) : P\int[V_I - V(a)]dF(a) > C\} \\ \mathcal{F}_O &= \{F(a) : P\int[V(a) - V_I]dF(a) > C\} \\ \mathcal{F}_A &= \{F(a) : |P\int[V(a) - V_I]dF(a)| \leq C\}. \end{aligned}$$

Accordingly, we can divide the citizens' population Ω into Ω_I, Ω_O, Ω_A, depending on their priors. If uninformed about the true opponent's type, citizens in the set Ω_I will vote for the incumbent, citizens in Ω_O will vote for the opponent and finally those in Ω_A will abstain. If informed about the opponent's type then prior beliefs clearly do not matter.

For our purposes it is important to distinguish group Ω_A from the rest. We can define citizens in this group as *strongly non-ideological* (although they can still be *weakly ideological*).

To compare different degrees of ideology we need a further simple definition:

Definition 2.3 *Assume citizens i and j have the same C and prior beliefs represented respectively by the distribution functions $F(a)$ and $G(a)$. Then we say that citizen i is more O-ideological (I-ideological) than citizen j if*

$$\int[V(a) - V_I]dF(a) > (<) \int[V(a) - V_I]dG(a)$$

When we don't want to distinguish between I-ideology and O-ideology we will simply say that agent i is more ideological than agent j.

It should be noted that in all the definitions of this section we consider a generic utility function $V(a)$, imposing on it no restriction other than that of being monotonic nondecreasing. The purpose of such definitions is to impose restrictions on the distribution functions rather than on the utility function. Given that we only want to characterize beliefs, independently of preferences, what is required on the functions $F(a)$ and $G(a)$ must be true for any nondecreasing utility function $V(a)$. Now notice that

$$\int [V(a) - V_I] dF(a) = \int V(a) dF(a) - V_I.$$

If we pose a further restriction and require Definition 2.3 to be valid for every nondecreasing function $V(a)$ then it is clear that the comparison of alternative distribution functions based on our definition of ideology is equivalent to using first order stochastic dominance.

$$\int [V(a) - V_I] dF(a) > (<) \int [V(a) - V_I] dG(a)$$

Under this more restrictive requirement we can introduce an indicator of ideology that will be useful in the rest of this section[11].

Definition 2.4 *Define π as an indicator of ideology s.t. an increase in π indicates an increase of O-ideology.*

Assumption 2.3 *Consider two distribution functions $F_{\pi_F}(a)$ and $G_{\pi_G}(a)$. Then $\pi_F > \pi_G$ if and only if $F_{\pi_F}(a) \leq G_{\pi_G}(a)$ $\forall a$.*

Thus, as π increases, agents become more Q-ideological or, alternatively, less I-ideological. Also, as $|\pi|$ increases we will say that agents become more ideological (without

[11] Defining ideological beliefs using first order stochastic dominance seems to make clearer the distinction between private interest and ideology. In a sense, an ideological belief must be independent of preferences (at least as long as we all agree on some basic premise, like that a is a valuable thing). If an individual believes that $F(a)$ stochastically dominates $G(a)$ then he would recommend $F(a)$ to every person with a nondecreasing utility function $V(a)$. This captures the difference between the fact that $F(a)$ is better for the ideological person and the fact that such person believes $F(a)$ to be better for everyone.

specifying). If we now define π_{-1}, π_0, and π_{+1} such that

$$\begin{aligned}
\int [V_I - V(a)] dF_{\pi_{-1}}(a) &= C \\
\int [V_I - V(a)] dF_{\pi_0}(a) &= 0 \\
\int [V(a) - V_I] dF_{\pi_{+1}}(a) &= C
\end{aligned}$$

then agents are defined as strongly I-ideological when they have $\pi < \pi_{-1}$, weakly I-ideological when $\pi_{-1} \leq \pi \leq \pi_0$, weakly O-ideological when $\pi_0 \leq \pi \leq \pi_{+1}$, and strongly O-ideological when $\pi > \pi_{+1}$. We will also refer to citizens for which $\pi_{-1} \leq \pi \leq \pi_{+1}$ as "abstainers".

Finally, it is useful to define the following sets, corresponding to possible realizations of the opponent's type:

$$\begin{aligned}
\mathcal{A}_I &= \{a : P[V(a) - V_I] < -C\} \\
\mathcal{A}_A &= \{a : -C \leq P[V(a) - V_I] \leq C\} \\
\mathcal{A}_O &= \{a : P[V(a) - V_I] > C\}
\end{aligned} \tag{2.9}$$

Figure 2-2 shows the partitioning of the opponent's type support in the case in which $V(a)$ is a linear function, and reports the critical values of π.

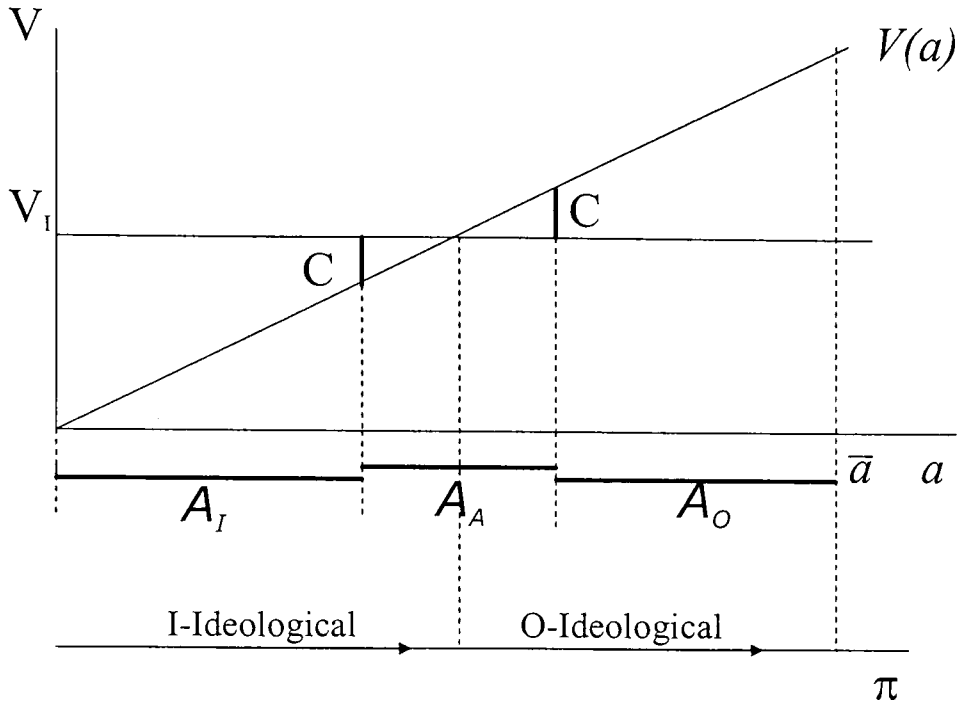
The value of information depends on the decision the citizen would make following only her priors. In particular, information is valuable in that it might change the decision taken when uninformed. Consider a strongly O-ideological agent. As π increases, the probability of realizations in $\mathcal{A}_I$ or $\mathcal{A}_A$ decreases, thus rendering the possibility of uninformed mistakes less likely. Therefore the value of information should decrease as π increases.

Things are more complex for weakly ideological agents and we need to introduce a further assumption in this case.

Assumption 2.4 $\pi_i > \pi_j \Rightarrow P \int_{\mathcal{A}_I \cup \mathcal{A}_O} |V(a) - V_I| [f_i(a) - f_j(a)] da \geq C \int_{\mathcal{A}_I \cup \mathcal{A}_O} [f_i(a) - f_j(a)] da$.

It is important to note that this assumption is at the same time both a restriction on the distribution functions considered, and a restriction on the possible partitions of

Figure 2-2: Partitioning the $\mathcal{A}$-space



the space $\mathcal{A}$. Now consider a weakly O-ideological agent, an agent that would abstain if uninformed. In this case an increase in π will clearly decrease the probability of events in $\mathcal{A}_I$; on the other side now the probability of events in $\mathcal{A}_O$ increases. We are left therefore with two opposite effects whose overall impact cannot be determined. Then Assumption 2.4 basically says that when an agent is O-ideological, as π increases we expect the increase of likelihood of events in $\mathcal{A}_O$ to dominate the corresponding reduction of likelihood of events in $\mathcal{A}_I$.

We can now state the following proposition.

Proposition 2.4 *Let us indicate with Δ_F and Δ_G the value of information corresponding respectively to π_F and π_G. Under Assumption 2.4 and for given E, M, w, r, C, we have that*

$$\begin{aligned}
1)\ \pi_G &< \pi_F < \pi_{-1} \Rightarrow \Delta_G < \Delta_F \\
2)\ \pi_{-1} &\leq \pi_G < \pi_F < \pi_0 \Rightarrow \Delta_G > \Delta_F \\
3)\ \pi_0 &\leq \pi_G < \pi_F \leq \pi_{+1} \Rightarrow \Delta_G < \Delta_F \\
4)\ \pi_{+1} &< \pi_G < \pi_F \Rightarrow \Delta_G > \Delta_F
\end{aligned}$$

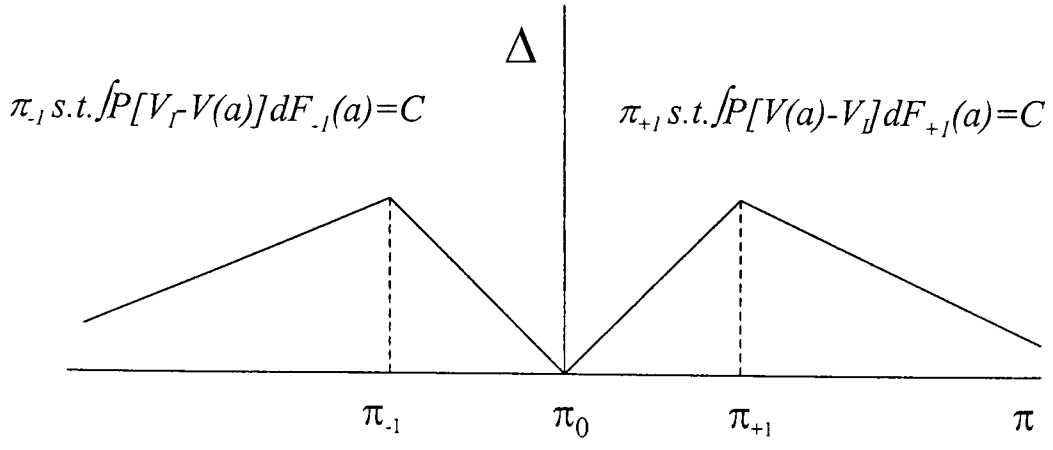
Proof. See Appendix 2A.

Following the foregoing discussion, the intuition for this result should be quite clear and has a simple representation in Fig. 2-3. Citizens that believe there is very little difference between the candidates (compared to the cost of voting) have little benefit from acquiring information: the expected utility from an informed versus an uninformed choice is very limited as not much difference is expected. Citizens who are extremely *independent* in their evaluation of candidates can therefore be better classified as *indifferent*: they will tend to attach little value to politics in general and therefore will remain generally uninformed. As priors become more and more ideological, the demand for information will increase, as the value of an informed decision increases too. The value of information reaches its maximum for those citizens that are exactly indifferent between voting or not: for such agents observing the realization of a carries a probability 1 of breaking the indifference. If we assume that indifferent[12] agents will abstain, there is a very high probability of a realization occurring in, for example, $\mathcal{A}_O$, thus making information extremely valuable. We will call the agents in a neighbourhood of this point *independent.* Starting from this maximum, the value of information will instead decrease monotonically for further increases in ideology. This happens when citizens's prior are strong enough to induce them to vote if uninformed: holding very strong priors means also to believe that it is not worth to acquire new information. We will refer to those agents as *partisan.*

It should be noted that Proposition 2.4 is stated for a given C. However, as C increases

[12]Note that the word "indifferent" here is used with a different meaning (indifference between voting and abstaining) with respect to the previously mentioned indifference between candidates.

Figure 2-3: The Value of Information



we should expect the number of uninformed agents to increase: citizens that, in spite of being sufficiently ideological, have a very high cost of voting (think for example of citizens living outside their home country) can be expected to remain rationally ignorant.

It is now possible to proceed, linking information and turnout.

Proposition 2.5

$\frac{\partial \Pr(T=1|Q)}{\partial Q} \geq 0$ *for abstainers and* $\frac{\partial \Pr(T=1|Q)}{\partial Q} < 0$ *for strongly ideological citizens.*

Proof. See Appendix 2A.

For abstainers the situation is analogous to that presented for a non-partisan polity in Proposition 2.2: information can only increase the probability of voting for citizens that would otherwise abstain with certainty. Things are just the opposite for partisans: information could lead them to discover that candidates are not as distant as they perceived, thus inducing them not to incur the cost of voting. Thus, the impact of information on turnout depends on ideology.

It is important, at this point, to understand what is the effect of ideology on turnout. A number of empirical studies tend to show that ideology matters for voting decisions[13]. This is quite intuitive and in line with what most political scientists would argue. Here, however, we found that ideology matters also for information acquisition and, in turn,

[13] See for example Palfrey and Poole (1987).

that information matters for turnout. What is the final effect of ideology on turnout then? Proposition 2.6 provides a result that takes into account the existence of both a direct and an indirect (via information acquisition) effect. Our conclusion is that the indirect effect is not enough to contradict the basic intuition that more ideological citizens are more prone to vote.

Proposition 2.6 $\pi_F > \pi_G > \pi_0 \Rightarrow Pr(O|F) > Pr(O|G)$; $\pi_F < \pi_G < \pi_0 \Rightarrow Pr(I|F) > Pr(I|G)$. *If the function* $|V(a) - V_I|$ *is symmetric around* 0 *then for any two prior distributions* $F(\cdot)$ *and* $G(\cdot)$ $|\pi_F| > |\pi_G| \Rightarrow Pr(T = 1|F) > Pr(T = 1|G)$

Proof. See Appendix 2A.

All the results presented in this and the previous section are given for a fixed cost of voting. The extensions of these results considering that C could be not constant across the population is straightforward and will be omitted. As long as C is independent of other characteristics, results for the whole population will not be affected.

Before moving to the empirical analysis, we can now spend a few words on the impact of information on collective choice. We proved that weakly non-ideological citizens increase their likelihood to vote when informed, while strongly ideological ones increase their likelihood to abstain. It is then clear that more information increases the chances to win of the better politician. It is therefore also worth attempting to spell out the implications of our results, although derived in a decision-theoretical context, for the literature on information aggregation that associates better information with an higher likelihood of turnout. Our model delivers such a link but also makes it conditional on voters' prior beliefs. Information is good also because it can induce abstention of otherwise uninformed extremists: this, clearly, makes no harm (and is actually beneficial) to the possibility of information aggregation to occur. At the same time, we can argue that the possibility of information aggregation in elections should be related to a number of individual (often observable) characteristics as well as to specific characteristics of the environment, mainly related to information supply by the mass media.

We are now ready to move to empirical investigation and test the predictions of our theoretical analysis for what concerns turnout. As a by-product, and although this is not the main task of this work, we will also be able to test most of the theoretical conclusions we reached about information acquisition.

2.5 The data

We will now proceed to verify the compatibility of our theoretical results with data. Empirical investigation will concern the 1997 general election in the United Kingdom[14]. Data on England, Scotland and Wales will be used; the political situation in Northern Ireland is substantially different from the rest of the country as the main cleavage is between the Catholic and Protestant populations rather than on the usual left-right dimension.

In the U.K. members of parliament (MPs) are elected one in each electoral constituency in a first past the post system. Since the executive needs the support of the parliament, the election is won by the party which obtains the larger number of MPs. Party leaders are candidates to become prime ministers; nevertheless, they need to win in their own constituency to be elected MP.

There are two major parties, Conservative and Labour, although other parties manage to win in some constituencies. In particular, the Liberal-Democratic party is well established nationally as the third force in the political arena. In Scotland, the independentist Scottish National Party is more than a third force, winning normally even more constituencies than the Conservatives. In the election we consider, the 1997 general election, the Labour party obtained a neat victory after 17 years of Conservative governments.

Our main source of data is the British General Election Study (BGES); this is a survey consisting of 3615 individual observations about voters who were interviewed a short time after the election took place. We had to drop a certain number of observations because some respondents did not answer all of the questions we used. For our purposes we will use in most cases a sample of 2769 observations.

The qualifying date for electoral registration expired several months before the election day[15]. We are interested in information acquisition and turnout of potential voters; clearly the motivations of non-registered voters are not captured in our model, at least if we think of the electoral campaign as the central moment to acquire information about candidates. Thus, only registered voters will be considered[16].

The first hurdle is to find a way to measure information. How can we measure political

[14]The election was held on the 1st of May 1997.

[15]That was 10th October 1996.

[16]This implies dropping a very limited number of observations. All estimations have been replicated including those observations and no relevant changes have been noticed.

knowledge? What is relevant for voting decisions? Is there any risk to bias the assessment with our priorities? Fortunately, as we noticed in Chapter 1, there is evidence of high correlations in the probabilities to be informed on any relevant political issues[17], and we can exploit the limited information we have with some confidence.

The BGES is particularly suited for the purpose of our analysis. Among other questions concerning the election, respondents received two sets of questions that can be used to establish how much they know about politics. In a first set of questions they were asked to write down as many candidates' names in their constituency as they could remember (with a maximum of 6). These names have then been checked and a point has been given for each correct answer. In a second set of questions, respondents received 7 statements on the British political and institutional system and were asked to say if they were true or false[18]. For each correct answer to these questions a score of 0.66 has been attributed to the agent[19]. The scores in the two set of questions have then been added up into a variable ($INFO$) that will be used as a measure of how much people know about British politics. This ranges from 0 to 10.62. An approximate graphical representation of the distribution of $INFO$ is reported in fig. 2-4. The continuous density function reported is normal with mean and variance of the observed $INFO$ (see tab. 1 in Appendix 2B)[20].

Another problem arises when measuring ideology. Disposition variables such as a person's interest in politics, sense of political efficacy etc. have been found to be quite important explanatory variables for electoral turnout. However, there are some reasons to be cautious about their usage. It is possible for example that the answer to these questions are respondents' rationalizations of their behaviour; also, it has been shown that responses sometimes are quite sensitive to the order of questions, which casts doubts on the validity of such indicators[21]. Nevertheless, to test some of our conclusions we need to rely on this type of information.

Two possible measures of partisanship are considered. One is the classical left-right self-placement, with zero being the extreme left and 10 the extreme right. We will trans-

[17] Delli Carpini and Keeter (1996).

[18] Statements are reported in Appendix B.

[19] The different weight is derived by Bayes rule: see Appendix B for details.

[20] It is possible to combine questions in different ways or to use only one of the two sets of questions to derive different indicators of political awareness. Such variations have limited impact on our results.

[21] See for example Bishop, Oldendick, and Tuchfarber (1984) and Abramson, Silver, and Anderson (1987)

form this variable by pulling together corresponding levels of extremism on both sides. This leads to a variable ($Left - Right1$) that assumes a value of zero if the original variable was 5, 1 if it was 4 and 6 etc. As a large number of respondents could not or did not want to place themselves in such scale, we end up with a smaller sample size. Some observations can be recovered by including at least a part of the non-respondents among the least ideological: this variable is then called $Left - Right2$. Further details can be found in Appendix 2B.

The second indicator ($Party$) measures instead how close respondents feel to their preferred party (if any) and has been built up by combining three questions. A full description of the variable $Party$ can be found in Appendix 2B.

The survey also includes information on a number of individual and household characteristics that can be used both as control variables or to test some specific predictions of our theory, and in particular those concerning information acquisition.

Household income is grouped, with 1 being the lowest and 16 the highest category. A few hundred of the interviewed refused to disclose information on income and this is the main reason for dropping part of the observations. The dataset does not contain instead any information on wage rates, that could be taken as a proxy for the opportunity cost of information gathering. We can use instead the number of hours spent on work ($Hours$). In spite of the common anecdotal evidence about reading newspapers at work (or, more recently, surfing the net), political information acquisition and time spent at work should not normally be considered as competitive ways of allocating time, at least for most people. If economic theory induces to think in those terms is probably because most of the literature on time allocation has concerned labour supply. In our case choosing how to allocate a given leisure time is probably a more appropriate description of the choice faced by agents: thus, the number of hours spent at work gives enough information in this sense. It should be clear that in this way we are not capturing a substitution effect but rather an endowment (of leisure) effect. Although it does not seem to appear in empirical literature on turnout, $Hours$ has also been used as an explanatory variable in the turnout equation to take into account the opportunity cost of voting on the day of the election, as this took place in a normal working day (Thursday).

The survey contains information on a number of socioeconomic characteristics that, in terms of our model, represent the parameters on each agent's production function of

information. These are education, sex, age, marital status and many more; while for some variables their link with the capability to acquire, process, and retain information seems quite obvious, other will mainly reflect the different networking possibilities of the various agents. Moreover they will be used as control variables in the turnout equation.

Information on the use of mass media has also been used. We know if the respondent regularly reads a newspaper and which one. In particular it is possible to distinguish between the regular readers of quality newspapers and the rest of the population. Details are reported in Appendix 2B. Information on canvassing and phone contacts between the interviewed and party representatives has also been used both to know whether respondents had information provided directly by parties and to infer about the effort of parties in different constituencies.

The BGES reports the electoral constituency of each observation. It is then possible to match this dataset with electoral results to measure the effect of election closeness on turnout probability[22]. Closeness is measured using the percentage difference between the winning candidate and the runner up in the constituency. This requires some kind of rational expectations assumption or, simply, the fact that people know about pre-electoral polls and that those polls are substantially correct. Other possibilities[23] have been considered instead, all giving the same results.

Information on the provenance of each observation has been used to match the BGES data with the Census (1991) data, to check for possible effects due to some relevant characteristics of the local environment, like unemployment rates, average education levels etc[24].

In general, our attempt will be to include all the variables that, for different reasons, have been considered by the empirical literature on turnout (see for example Matsusaka and Palda, 1999). For this reason the list of variables is quite long, and the standard errors are often high because of multicollinearity. However, this strategy should lead to robust results for what concerns our variables of interest.

[22] Data on electoral results are taken from Boothroyd (2002).

[23] Instead of using the results of the current election (using then a rational expectation argument), it is possible in principle to use past elections. One serious limitation is in this case represented by the fact that the boundaries of most constituencies were changed between 1992 (the year of the previous general election) and 1997. Also, it can be argued that constituencies' size matters for the probability to be pivotal and therefore absolute and not percentage differences should be used. We have tried these different alternatives and the results are not sensitive to the changes.

[24] Data from the Census are at the level of Districts, local administration units reported in the BGES.

Finally, to test for potential endogeneity of information in the turnout equation, we will use four instrumental variables that are assumed to affect information acquisition but not turnout directly. We will introduce them in the next section.

Data description and summary statistics are reported in Appendix 2B.

2.6 Estimation strategy and empirical specification

2.6.1 Information and turnout without ideology

A number of testable predictions were derived in section 2.3, concerning information acquisition and the positive link between information and turnout. This sub-section illustrates the estimation strategy to test those results, with a special focus on the role of information as an explanatory variable for turnout; we would like to be able to establish a *causal* link between the two, and for this purpose we need to deal with potential endogeneity problems.

We can start by defining a citizens' utility from voting as $P|\int[V(a)-V_I]f(a)da|-C$, where $f(a)$ can be a degenerate distribution in the case of an informed voter. For the purpose of empirical investigation we will also consider non-instrumental voting, including the benefit derived from fulfilling a civic duty D to define the variable

$$B_0 = P|\int[V(a)-V_I]f(a)da| + D - C. \tag{2.10}$$

B_0 is a latent (unobservable) variable and turnout T is a binary indicator such that

$$\begin{aligned} T &= 1 \text{ if } B_0 > 0 \\ T &= 0 \text{ if } B_0 \lesssim 0 \end{aligned}$$

We can approximate B_0 by using a linear random utility model:

$$B_0 = \beta' \mathbf{X} + \varepsilon \tag{2.11}$$

where $\mathbf{X}$ is a vector of characteristics of the individual considered and of the environment[25]

[25] This includes a vector of 1s, and therefore β' includes a constant term.

(including P) and ε is a white noise disturbance including the non-instrumental benefit D (some imperfect indicators of the sense of civic duty can however be included in $\mathbf{X}$). We can then say that

$$\begin{aligned} \Pr[T &= 1|\mathbf{X}] = \Pr[B_0 > 0|\mathbf{X}] \\ &= \Pr[\beta'\mathbf{X} + \varepsilon > 0] \\ &= \Pr[\varepsilon < \beta'\mathbf{X}] = \mathbf{F}(\beta'\mathbf{X}) \end{aligned} \tag{2.12}$$

Appropriate assumptions over the distribution $\mathbf{F}(\cdot)$ will allow estimation of the (2.12) by maximum likelihood. We will assume $\mathbf{F}(\cdot)$ is the logistic distribution function and therefore we will estimate $\Pr[T = 1|\mathbf{X}]$ by maximum likelihood logit.

In most empirical literature turnout is estimated using some analogous procedure. We will start by using our data to estimate equation (2.12), including all the variables that have traditionally been identified as relevant. Results are reported in tab. 2.3 and will be discussed in the next section.

Let us now introduce political information and indicate by $\widehat{q}$ the realization of the random variable q after t and k have been acquired and before voting We can then say that

$$\begin{aligned} T &= 1 \text{ if } B_0 > 0 \text{ and } \widehat{q} = 1 \\ T &= 0 \text{ if } \{B_0 > 0 \text{ and } \widehat{q} = 0\} \text{ or } B_0 \leq 0 \end{aligned}$$

For simplicity we will define a new latent variable $B_1(B_0, \widehat{q})$ and choose a linear representation of the form

$$B_1 = \beta_1'\mathbf{X} + \beta_2\widehat{q} + \varepsilon \tag{2.13}$$

We then have that

$$\Pr[T = 1|INFO, \mathbf{X}] = \mathbf{F}(\beta_1 INFO + \beta_2'\mathbf{X}) \tag{2.14}$$

where we replaced $\widehat{q}$ with its observable counterpart $INFO$. Estimating the (2.14) is a correct procedure only if information acquisition is orthogonal to turnout. In Feddersen and Pesendorfer (1996), for example, people are randomly informed or uninformed about the true state of the world. However, the benefit D in equation (2.1) can be an important

motivation for voting, in the same way b in (2.4) is for information acquisition. The two types of psychic benefit are very likely to be correlated.

Therefore information could be an endogenous explanatory variable and the coefficient estimates of (2.14) biased. To overcome this potential problem we will then estimate the following triangular system, where i refers to a generic observation (citizen) in our sample:

$$\begin{aligned} INFO_i &= \alpha_1^{'}\mathbf{X}_i + \alpha_2^{'}\mathbf{Z}_i + u_{1i} && (2.15) \\ B_1 &= \beta_1 INFO_i + \beta_2^{'}\mathbf{X}_i + u_{2i} && (2.16) \\ T_i &= 1 \text{ if } B_1 > 0 \\ T_i &= 0 \text{ if } B_1 \leq 0 \end{aligned}$$

where $\mathbf{X}$ is again a vector of covariates representing both individual and constituency characteristics and assumed to affect both turnout and information. Our identifying covariates are represented by the vector $\mathbf{Z}$: these explanatory variables are assumed to affect political knowledge but not directly the turnout decision.

It is clear that if this is the structural model, then simple probit estimates of (2.16) will suffer of endogeneity bias as the two error terms u_{1i} and u_{2i} are correlated . By using instrumental variables we should also be able to assess the relevance of this bias.

Treating $INFO$ as a continuous variable, the system is estimated in two steps. Equation (2.15) is a reduced form containing all the exogenous covariates of our model. The first step consists of estimating the reduced form (2.15) by OLS and get the residuals $\widehat{u}_{1i} = INFO_i - \widehat{\alpha}_1^{'}\mathbf{X}_i - \widehat{\alpha}_2^{'}\mathbf{Z}_i$.

We can then estimate the equation

$$T_i = \beta_1 INFO_i + \beta_2^{'}\mathbf{X}_i + \beta_3\widehat{u}_1 + u_{2i} \qquad (2.17)$$

by logit maximum likelihood. This provides both consistent (though not efficient) estimates of $(\beta_1, \beta_2^{'})$, as well as an endogeneity test: if β_3 is insignificant we cannot reject the null hypothesis that $INFO$ is weakly exogenous in the turnout equation.

The vector $\mathbf{Z}$ is composed of four variables that are assumed to influence information acquisition but not directly turnout. The variable *Salience* attempts at capturing the

salience on media of each constituency during the electoral campaign[26]. For this purpose information from a major national newspaper, The Guardian, has been used. *Salience* consists of a dummy equal to 1 if an article focusing on the electoral contest of a specific constituency appeared on this newspaper during the campaign. There is no specific reason for using The Guardian apart from the fact that it is a national quality newspaper and its archive is easily accessible: any newspaper with the same characteristics could be used instead, the only purpose being to capture salience (not just on newspapers). Our assumption is that people living in more salient constituencies are more exposed to political information and therefore, for a given effort in news-gathering, will know more about politics in the day of the election.

A second instrument is *bbc100*. This is a dummy equal to 1 for citizens living in constituencies on which the BBC decided to focus its attention on the night of the election (these are listed on the BBC web page): they were the expected closest 100 Consevative-held constituencies (and therefore the decisive ones as the Conservative were commonly expected to loose constituencies). They were described on the BBC web-page as "the battleground" of the election: we expect these constituencies to get larger media coverage during the electoral campaign.

A third instrument is represented by a dummy variable (*big-shot*) equal to 1 when a nationally relevant politician was candidate in the constituency. We define as big-shots all the current and past members of cabinet, the members of the Labour "shadow-cabinet" and the leader of the third major party (Liberal-Democratic), Paddy Ashdown[27]. Finally, we include an instrument on media usage: a dummy equal to 1 if the agent reads regularly a quality newspaper.

It is known that the estimated standard errors from this method are not correct. However, Monte Carlo evidence tend to show that the asymptotically correct standard errors are no more effective in large finite samples than the conditional standard errors (see Guilkey, Mroz, and Taylor, 1992). Nevertheless, to overcome any potential problems, standard errors of relevant variables have been estimated by bootstrap (with 1000

[26] This is defined as the last 30 days before the election day.

[27] A "shadow-cabinet" is put in place by the opposition party and is composed by those who, in case of victory, most likely will become ministers. A shadow foreign secretary, for example, follows very closely the government foreign policy and is supposed to be able to control and propose alternatives. This makes shadow-cabinet members quite popular on the media.

repetitions).

Associated with this two-step logit regression model there is an endogeneity test to determine whether there is overlapping between the set of unobservables affecting equations (2.15) and (2.16). It consists of a simple t-test for significance of the coefficient β_3 of the estimated error term.

Finally we will test the validity of the instruments. This can be done in several different ways. One possibility is to compare a logit regression of turnout on all exogenous variables and instruments (unrestricted model, first column of Tab. 2.6) with the same regression where instruments are excluded but fitted values from the first stage regression are included (restricted model, fourth column of Tab. 2.4); ideally, we would like the two to be not "too different": we can then perform a chi-square test based on the likelihood function. Another possible method will be discussed later, when results are presented.

2.6.2 Information, turnout, and ideology

The first step in analysing the role of ideology is to estimate the information function in order to test proposition 2.4. As noticed in the previous sub-section, estimating the (2.15) is interesting for the sake of understanding the determinants of political knowledge and testing our proposition 2.1. Ideology is considered by estimating the equation

$$INFO_i = \alpha_1^{'}\mathbf{X}_i + \alpha_2^{'}\mathbf{Z}_i + \alpha_3 ID_i + u_i \tag{2.18}$$

where ID represents one of the three measures of ideology introduced in the previous section ($Left-right1$, $Left-right2$, and $Party$). Suppose there are K types of citizens ranked according to their degree of ideological motivation. Then ID is a categorical variable and we will introduce $K-1$ dummies in the regression. We expect to find a non-monotonic pattern in such dummies, where estimated parameters should first increase with ideology and then decrease. Estimation is by OLS.

We can finally turn to the impact of ideology on turnout, and in particular to how the effects of information on turnout differ according to ideology. This will be done by

estimating the equation

$$\begin{aligned} B_2 &= \beta_1 INFO_i + \beta_2' \mathbf{X}_i + \beta_3 ID_i + \beta_4 ID_i \times INFO_i + u_{2i} \qquad (2.19) \\ T_i &= 1 \text{ if } B_2 > 0 \\ T_i &= 0 \text{ if } B_2 \leq 0, \ i = 1, ..., N. \end{aligned}$$

Differently from the (2.18), now ID is treated as a continuous variable, in order to interact it with $INFO$. Indicating with $\overline{INFO}$ the average of $INFO$, we expect $\beta_3 + \beta_4 \times \overline{INFO} \geq 0$ (from proposition 2.6), and $\beta_4 \leq 0$ (from proposition 2.5).

2.7 Results

We start by running a logit regression of turnout on a set of variables that both theoretical and empirical literature have identified as relevant. Estimations of (2.12) are reported in Table 2.3. In column 2 income, education and churchgoer are considered as categorical variables, in column 1 they are numerical variables (thus we impose a linear restriction). Although we can accept such a restriction for any of these variables in isolation, this is not true for the three together, as a comparison of the log-likelihood scores would formally show. Therefore in the subsequent analysis we will only consider the case where fixed effects for all three variables are included. However, all results do not change in any substantial respect.

As previously explained, our list of explanatory variables is as comprehensive as possible. This means most variables are correlated, implying relatively high standard errors. The sign of coefficients show some surprises if compared with most previous findings. The most significant explanatory variables (at 5% significance level and above in both columns) are *voted92, married, hours, canvasser, churchgoer,* and *income. Education* significance level is just below 5%. More surprisingly, *age* and *sex* seem to be completely irrelevant. The fact that past voting behaviour is an extremely important explanatory variable reinforces the idea that there are relevant individual-specific unobservables in driving turnout behaviour. However, this variable is also correlated with other explanatory variables. Column 3 in Table 2.3 presents results for a slightly more parsimonious model where *voted92* and *hours* (a variable not normally considered in the voting literature) have been

excluded: these estimates have a decidedly more "traditional" flavour with *age* below the 5% significance level and *sex* much closer to it. Our data show therefore no real surprises in this sense.

It should also be noted the result on the marginality of the electoral constituency, an issue that has received careful attention in the literature, with quite controversial results. We find that the closeness of the election has the expected sign but also that its significance is definitely too low to be considered of any relevance[28]. Other socioeconomic characteristics of constituencies also do not have any significant impact; once again, most of the effect at the constituency level is captured by aggregate turnout. The weekly number of hours spent at work proves to be an important explanatory variable, although being neglected so far by empirical research.

In table 2.4 we report estimates of the turnout equation when we include information. Looking at the first two columns[29], it is clear that *INFO* is a good predictor of turnout; both its magnitude and significance level seem to suggest that information is amongst the most important explanatory variables for voters' electoral participation. This result is very robust to variations in the specification adopted. Education now becomes completely insignificant, because of an obviously high correlation with $INFO$ and suggesting that the effect of education on turnout is mainly driven by information. Age has a negative and convex effect (although still not significant at 5% level). This could suggest that the positive effect of age that some studies seem to find could be due to the larger experience and knowledge of political matters that older people might have accumulated during their lives: apart from this (and remembering that we also control for the number of hours spent at work) age tends to have a negative impact on voting, as one would expect considering that the elderly are on average less fit and therefore have an higher cost of going to the poll station.

We still need to address the potential endogeneity problem that might occur when regressing turnout on information. Results of the first stage regressions (2.15) are shown in Table 2.5. These regressions are also of interest for their own sake, as they can be seen as estimates of a political-knowledge function. Since our main equation is the (2.16), we are

[28] Although its significance raises considerably when we drop the aggregate turnout variable.

[29] As before, the only difference between the two is that in the second some variables are treated as categorical.

not making any attempt to have a correct specification of the demand for information. Equation (2.15) is then just a reduced form that makes use of all available exogenous variables: this can affect efficiency but not consistency of estimates.

The first thing to note is that our instruments are significant and show the expected sign. They are clearly correlated as they try to capture similar effects and this makes their individual significance even more remarkable. Among other exogenous variables, both *education* and *income* are very strong predictors of political knowledge; it should not be overlooked the fact that we get this result in spite of controlling for the most important covariates that are normally used to explain income in itself. It is then possible to conclude quite safely that political information can be treated as a normal good.

Age and *sex* have very strong effects; the first probably because, as we said, more experienced citizens have attained a larger "stock" of political knowledge, the second reflecting different networking possibilities some time faced by members of the two sexes, as well as different forms of socialising in general. The length of residence in a given constituency and union membership could also capture experience and networking effects; in the case of trade unions, they often spend a remarkable effort in informing their members about political matters, especially related to labour policy.

There is significant correlation between information and the number of hours devoted to work. Following the discussion on this point in the previous section, it is reasonable to assume that, being the leisure time of full time workers lower, the opportunity cost of time devoted to information gathering is higher, as confirmed by the sign of the coefficient.

As a general comment on individual-level variables, it is possible to conclude that the personal technology used in receiving, processing and retaining news plays a crucial role in information acquisition. These are the parameters that in the model we indicated by E and whose signs are very well predicted by our theory: they all appear to be significant predictors of political knowledge.

Constituency-level variables, differently from what happens in the turnout equation, matter for information. First of all voters are substantially more informed in constituencies with closer competitions. The effect of closeness on political knowledge is definitely stronger than that on turnout (compare Tab. 2.5 with Tab. 2.3). This could simply be due to the fact that politicians and parties put more effort in marginal constituencies (as suggested for example in Aldrich, 1993, and Shachar and Nalebuff, 1999). However, we

control for this by using a measure of party effort in each constituency. As an alternative interpretation, it is possible that the individual demand for political information increases when the probability to be a pivotal voter is higher: this would provide evidence of another form of voters' rational behaviour. Such conjecture will be then investigated more closely in Chapter 4.

The effect of education at the constituency level is rather unclear, with the percentage of higher degrees having a positive effect and that of degrees a negative one (although less sizeable and less significant). Quite surprisingly unemployment rate increases political awareness.

Voters are also better informed in constituencies that received a more extensive newspaper coverage during the electoral campaign, as shown by *salience;* mass media seem to be quite effective in improving the knowledge citizens have about political matters, as shown also by the sizeable effect of being a quality-paper reader as well as by *bbc100.*

Let us turn to the endogeneity issue now. For this purpose we run a logit regression of turnout including among the covariates both observed information and fitted residuals from the first stage regression. In the third column of Tab. 2.4 it is possible to see that the sign of *INFO* is unchanged and its magnitude much larger. Although the z-statistic is now substantially lower, information is still comfortably significant at the 5% level. However, even more importantly, residuals are not significant; thus, on the basis of this evidence, we cannot reject the null hypothesis that information is weakly exogenous in the turnout equation. The overidentification test presented in the previous section is easily passed by our instruments: twice the difference between the log-likelihood of the equation in the first column of Table 2.6 and the equation in the fourth column of Table 2.4 is 1 and a chi-test is passed very comfortably. To this we can add a further test, reported in the second column of Table 2.6. If *INFO* is exogenous then the validity of instruments can be checked by including in the turnout equation both *INFO* and the four instrumental variables. If instruments are valid then the log-likelihood of this equation should be not too different from the restricted model when the four instruments are excluded. Comparing column two in Table 2.6 with column two in Table 2.4, we find that the log-likelihood ratio statistic for this restriction is equal to 4.18. With four degrees of freedom this test statistic is not significant at the 30% level.

Thus, from Table 2.6 we can conclude that the endogeneity test reported in the third

column of Tab. 2.4 is valid. Although, as one would expect, it is clear from the first stage regression that there are several variables driving both information and turnout, nevertheless we can safely assume that none of them has been omitted and therefore we can refer to the estimates of Tab. 2.4 as substantially correct. This will also allow us to proceed in further estimations ignoring the endogeneity issue.

We can now analyse the impact of ideology on information acquisition. Estimates of the (2.18) are reported in table 2.7. They support the predictions of our model. In the case of left-right self-placement, both *Left-right1* and *Left-right2* deliver similar conclusions. As compared with the omitted types (the least ideological), political knowledge first increases, reaching its peak (both in parameter size and significance) at the third category, then decreases and becomes insignificant for the last category: thus, the most ideological types are not significantly different from the least ones. When using party identification results are very similar. The peak is now in the fourth (out of five) category and both the second and the fifth ones are not significantly different from the first one.

Coming to turnout, our theory shows overall a good compatibility with data analysis. Both *INFO* and all our measures of ideology have the expected signs and are significant. First of all, the sign of $\beta_3 + \beta_4 \times \overline{INFO}$ is positive in all cases, which confirms once more the important role of ideology in fostering turnout . Remarkably, the interaction term has always a negative sign, although it is significant only when we use party identification as an indicator of ideology. Party identification, however, is also the variable that shows a larger added value to our regression, as shown by the Pseudo-R^2 statistic. It seems clear that our theory of the role of ideology in elections fits the case of party identification much better than that of left-right self-placement. This, overall, seems to provide evidence of an indirect effect of ideology (in the form of party identification) that, as we have seen in theory, can push towards a reduced turnout when information in received. The role of information is clearly more subtle here than before. Other things equal, information increases turnout on average even when we control for ideology (the sign of $\beta_2 + \beta_4 \times \overline{ID}$ is always positive); this, however, is not true any more for extremists. Take for example an agent with $Party\ Identification = 5$. In this case the marginal effect of information is $0.609 - 0.129 \times 5 = -0.045$. More in general, it seems clear that the positive impact of information on turnout tends to vanish (and, in the limit, to be reverted) with increasing ideology, accordingly with our theoretical predictions.

2.8 Conclusion

This chapter analyses the interactions between ideology, political information acquisition and electoral turnout and provides empirical evidence about their links. Information acquisition is modelled as an individual production function: citizens "produce" their own information by using mass media and time. Voters are endowed with different technologies, reflecting their ability to acquire, process and retain information. The parameters that determine different productivities are then represented by a series of individual characteristics like education, income, age etc. as well as by the supply of information, in the form of mass media coverage of political issues. This theoretical analysis leads to testable propositions about the links between individual and environmental characteristics and political knowledge.

The demand for political information also depends on ideological prior beliefs on candidates. In particular, it is possible to show that the least informed citizens are those with the weakest and the strongest ideological beliefs. In the first case, agents are so indifferent between candidates that the expected benefit of acquiring information does not cover its costs: contrarily to common wisdom, extremely "independent" citizens can be far from the ideal that a participative vision of democracy would require. At the same time, people with extreme prior beliefs will be confident enough in their opinions and again will not (ex ante) find useful to acquire information. Thus, we expect the most informed citizens to be slightly partizan: a moderate amount of ideology can therefore be useful to the functioning of democratic systems.

We then link ideology and information to turnout. While information has a positive effect on the likelihood of voting of non-ideological agents, it has instead a negative effect on the ideologized. Since those with strong priors are more likely to vote in the "wrong" way, this result confirms the importance of information for good collective decision-making and public officials accountability.

Our theory is capable of explaining most typical results of empirical research, like the positive effect of education on turnout. Moreover, through the interaction between ideology and information acquisition, we can derive new predictions: of particular relevance is the fact that information should have a positive impact on turnout only for non-partisan voters.

Empirical evidence on the 1997 general election in the United Kingdom is provided. Using a number of questions about candidate names and British politics in general, we can build up a measure of political knowledge that can be used to analyse the information-turnout relationship. Information is one of the most relevant and robust predictors of turnout. We estimate this relationship using both a simple logit and a two-step instrumental variables logit: in both cases the idea that political information is relevant for turnout seems well supported. More generally, we can safely conclude that our theoretical model shows a high degree of compatibility with data. As a by-product of this analysis we are also able to assess the role of individual and environmental characteristics on political awareness. From our estimates it is also clear that mass media are very important in determining political knowledge and, through this channel, electoral turnout.

This analysis has consequences for the way to think of the role of information and mass media in democratic systems. Overall, our findings show that information matters for electoral behaviour, thus contradicting the "behaviour irrelevance" hypothesis. There is instead some compatibility with the "outcome irrelevance" hypothesis. However, by taking information as exogenous, models leading to full information equivalence neglect incentives to acquire information. We show that, instead, in a very polarized polity, little information acquisition will occur, and ideology rather than information could determine policy outcomes. The same can happen to an extremely non-polarized population.

A consolidated research shows the importance of information in agency relations; a more recent and fast growing literature consistently finds evidence of a link between public officials' performance and information availability. By unveiling the impact of information on turnout, this chapter shows a possible rationale for politicians' responsiveness to an informed public opinion. We can conclude that mass media and voters' personal resources play a crucial role in democratic decision-making: formally democratic institutions might be emptied of their substantial content if good political information is either unavailable or beyond most voters' reach.

2.9 Appendix 2A: proof of results

Proof of Lemma 2.1 $\Delta = \int [W^*(a) - \widetilde{W}] f(a)da.$

Remember that

$$W^*(a) = \max_{\{T\}} T(P|V(a) - V_I| - C)$$

and define

$$W^* = \max\{0, P \int |V(a) - V_I| f(a)da - C\}$$

Also

$$\widetilde{W} = \max_{\{T\}} TP| \int [V(a) - V_I] f(a)da| - C)$$

which means

$$\widetilde{W} = \max\{0, P| \int [V(a) - V_I] f(a)da| - C\}$$

For Δ to be positive it is sufficient to prove that

$$\max\{0, \int |V(a) - V_I| f(a)da - C, 0\} \geq \max\{0, | \int [V(a) - V_I] f(a)da| - C\}$$

If we define

$$\begin{aligned} \mathcal{A}_- &= \{a : [V(a) - V_I] < 0\} \\ \mathcal{A}_+ &= \{a : [V(a) - V_I] \geq 0\} \end{aligned}$$

then it is clear that

$$\begin{aligned} \int |V(a) - V_I| f(a)da &= \int_{\mathcal{A}_+} [V(a) - V_I] f(a)da + \int_{\mathcal{A}_-} [V_I - V(a)] f(a)da \\ | \int V(a) - V_I f(a)da| &= | \int_{\mathcal{A}_+} [V(a) - V_I] f(a)da - \int_{\mathcal{A}_-} [V_I - V(a)] f(a)da| \end{aligned}$$

from which

$$\int |V(a) - V_I| f(a)da - C \geq | \int [V(a) - V_I] f(a)da| - C \qquad (2A.1)$$

If $P \int |V(a) - V_I| f(a)da \leq C$ then $W^* = 0$. But then 2A.1 implies that $| \int [V(a) - V_I] f(a)da| \leq C$ and therefore $\widetilde{W} = 0$.■

Proof of Proposition 2.2 Let us indicate with $\widehat{q} \in \{0,1\}$ the fact of being ex post informed ($\widehat{q}=1$) or not ($\widehat{q}=0$). For an uninformed citizen we have

$$Pr(T=1:\widehat{q}=0)=0$$

while for an (ex post) informed citizen, the probability to vote (*ex ante*) is

$$\begin{aligned} Pr(T &= 1:\widehat{q}=1) = \Pr(a:|[V(a)-V_I]|-C>0) \\ &= \int_{\mathcal{A}_I} dF(a) + \int_{\mathcal{A}_O} dF(a) \geq 0 \end{aligned}$$

where $\mathcal{A}_I$ and $\mathcal{A}_O$ are the sets defined in (2.9).

If $P[V(\overline{a})-V_I]>C$ and $P[V_I-V(0)]>C$ then $Pr(T=1|\widehat{q}=1)>0$.

The probability to vote is then given by the probability to be informed multiplied by the probability to vote when informed, i.e.

$$Pr(T=1|Q)=Q\Pr(T=1|\widehat{q}=1) \tag{2A.2}$$

from which the result follows immediately.■

Proof of Proposition 2.3 From the 2A.2 we have that

$$Pr(T=1|E,M,w,r)=Q(E,M,w,r)Pr(T=1|\widehat{q}=1)$$

We also know from Proposition 2.1 that

$$\frac{\partial Q(E,M,w,r)}{\partial E}\geq 0$$

from which it follows that

$$\frac{Pr(T=1|E,M,w,r)}{\partial E}=\frac{\partial Q(E,M,w,r)}{\partial E}\times Pr(T=1|\widehat{q}=1)\geq 0$$

Similarly we can prove the rest of the proposition.■

Proof of Proposition 2.4 Let us focus on the positive part of the diagram in Figure 2.3. Cases 3) and 4) refer respectively to weak and strong O-ideologies. Analysis will apply analogously to cases 1) and 2) (respectively strong and weak I-ideologies). Consider first a weakly O-ideological citizen. The value of information in such case is given by the probability information will induce a switch to a vote for I plus the probability it will induce a vote for O, i.e.

$$\Delta = \int_{\mathcal{A}_I} (P[V_I - V(a)] - C)dF(a) + \int_{\mathcal{A}_O} (P[V(a) - V_I] - C)dF(a)$$

Given two distributions F and G we want to prove that $\pi_F > \pi_G \Rightarrow \Delta_F > \Delta_G$ i.e.

$$\begin{aligned} &\int_{\mathcal{A}_I} (P[V_I - V(a)] - C)dF(a) - \int_{\mathcal{A}_I} (P[V_I - V(a)] - C)dG(a) + \\ &\int_{\mathcal{A}_O} (P[V(a) - V_I] - C)dF(a) - \int_{\mathcal{A}_O} (P[V(a) - V_I] - C)dG(a) \quad (2A.3) \\ > \; & 0 \end{aligned}$$

Define $s(a) = [V(a) - V_I]$. Assumption 2.4 implies

$$\begin{aligned} &- \int_{\mathcal{A}_I} Ps(a)[f(a) - g(a)]da - \int_{\mathcal{A}_I} C[f(a) - g(a)]da \\ &+ \int_{\mathcal{A}_O} Ps(a)[f(a) - g(a)]da - \int_{\mathcal{A}_O} C[f(a) - g(a)]da \\ > \; & 0 \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} &- \int_{\mathcal{A}_I} Ps(a)dF(a) - \int_{\mathcal{A}_I} CdF(a) \\ &+ \int_{\mathcal{A}_I} Ps(a)dG(a) + \int_{\mathcal{A}_I} CdG(a) + \\ &\int_{\mathcal{A}_O} Ps(a)dF(a) - \int_{\mathcal{A}_O} CdF(a) \\ &- \int_{\mathcal{A}_O} Ps(a)dG(a) + \int_{\mathcal{A}_O} CdG(a) \\ > \; & 0 \end{aligned}$$

$\Rightarrow 2A.3.$

Now consider a strongly O-ideological citizen. The value of information is in this case given by:

$$\Delta = \int_{\mathcal{A}_I} 2P[V_I - V(a)]dF(a) + \int_{\mathcal{A}_A} (P[V_I - V(a)] + C)dF(a)$$

i.e. the value due to a potential shift to a change in favour of I plus the value due to a shift in favour of abstention. Now we want to prove that $\pi_F > \pi_G \Rightarrow \Delta_F < \Delta_G$ i.e.

$$\begin{aligned} & \int_{\mathcal{A}_I} 2P[V_I - V(a)]dF(a) + \int_{\mathcal{A}_A} (P[V_I - V(a)] + C)dF(a) - \\ & - \int_{\mathcal{A}_I} 2P[V_I - V(a)])dG(a) - \int_{\mathcal{A}_A} (P[V_I - V(a)] + C)dG(a) \qquad (2A.4) \\ < \; & 0 \end{aligned}$$

It will be useful to adopt the following notation:

$$\begin{aligned} \mathcal{A}_I &= [\underline{a}, \widehat{a}] \\ \mathcal{A}_A &= [\widehat{a}, \widehat{\widehat{a}}] \\ \mathcal{A}_O &= [\widehat{\widehat{a}}, \overline{a}] \end{aligned}$$

Integrating the 2A.4 by parts we get:

$$\begin{aligned} & 2P[V_I - V(\widehat{a})]F(\widehat{a}) - 2P[V_I - V(\underline{a})]F(\underline{a}) + \int_{\mathcal{A}_I} 2PV'(a)F(a)da + \\ & +(P[V_I - V(\widehat{\widehat{a}})] + C)F(\widehat{\widehat{a}}) - (P[V_I - V(\widehat{a})] + C)F(\widehat{a}) + \int_{\mathcal{A}_A} PV'(a)F(a)da - \\ & -2P[V_I - V(\widehat{a})]G(\widehat{a}) + 2P[V_I - V(\underline{a})]G(\underline{a}) - \int_{\mathcal{A}_I} 2PV'(a)G(a)da - \\ & -(P[V_I - V(\widehat{\widehat{a}})] + C)G(\widehat{\widehat{a}}) + (P[V_I - V(\widehat{a})] + C)G(\widehat{a}) - \int_{\mathcal{A}_A} PV'(a)G(a)da \end{aligned}$$

Now notice that

$$\begin{aligned} & 2P[V_I - V(\underline{a})]F(\underline{a}) = 2P[V_I - V(\underline{a})]G(\underline{a}) = 0 \\ & P[V_I - V(\widehat{a})] = C \\ & P[V_I - V(\widehat{\widehat{a}})] = -C. \end{aligned}$$

We are left with

$$
\begin{aligned}
&2CF(\widehat{a}) + \int_{\mathcal{A}_I} 2PV^{'}(a)F(a)da \\
&-(C+C)F(\widehat{a}) + \int_{\mathcal{A}_A} PV^{'}(a)F(a)da - \\
&-2CG(\widehat{a}) - \int_{\mathcal{A}_I} 2PV^{'}(a)G(a)da - \\
&+(C+C)G(\widehat{a}) - \int_{\mathcal{A}_A} PV^{'}(a)G(a)da
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\Delta_F - \Delta_G \quad = \quad & \int_{\mathcal{A}_I} 2PV^{'}(a)[F(a) - G(a)]da + \\
& \int_{\mathcal{A}_A} PV^{'}(a)[F(a) - G(a)]da
\end{aligned}
$$

But $F(a) \leq G(a) \forall a$ which implies $\Delta_F \leq \Delta_G$.■

Proof of Proposition 2.5 The proof in the case of weakly-ideological citizens proceeds along the lines of the proof of proposition 2.2.

When agents are strongly-ideological we have, for uninformed citizens

$$
Pr(T = 1|\widehat{q} = 0) = 1
$$

while for an *(ex post)* informed citizen, the probability (*ex ante*) to vote is

$$
\begin{aligned}
1 \quad > \quad & \Pr(T = 1|\widehat{q} = 1) = \Pr(a|a \in \mathcal{A}_I \cup \mathcal{A}_O) = \\
= \quad & \int_{\mathcal{A}_I \cup \mathcal{A}_O} dF(a) > 0
\end{aligned}
$$

Note that the probability to vote conditional on being informed is the same both for strongly and weakly ideological citizens.

The probability to vote is then given by the probability to be informed multiplied by

the probability to vote when informed, i.e.

$$\begin{aligned} Pr(T &= 1|Q) = Q\Pr(T = 1|\widehat{q} = 1) + (1 - Q)Pr(T = 1|\widehat{q} = 0) \\ &= 1 - Q(1 - \Pr(T = 1|\widehat{q} = 1)) \end{aligned}$$

Proposition 2.5 follows from the fact that $\Pr(T = 1|\widehat{q} = 1) < 1$.■

Proof of Proposition 2.6 The probability of voting under the distribution function F is:

$$Pr(T = 1|F) = Q_F Pr(T = 1|\widehat{q} = 1) + (1 - Q_F)Pr(T = 1|\widehat{q} = 0))$$

Consider two weakly O-ideological distributions F and G s.t. $\pi_F > \pi_G$. Then $\Delta_F > \Delta_G$ and $Q_F > Q_G$. Thus

$$Pr(T = 1|F) = Q_F Pr(T = 1|\widehat{q} = 1) > Q_G Pr(T = 1|\widehat{q} = 1) = Pr(T = 1|G)$$

If instead $F, G \in \mathcal{F}_O$ then

$$\begin{aligned} Pr(T &= 1|F) = 1 - Q_F[1 - Pr(T = 1|\widehat{q} = 1)] \\ Pr(T &= 1|G) = 1 - Q_G[1 - Pr(T = 1|\widehat{q} = 1)] \end{aligned}$$

Now $\pi_F > \pi_G \Rightarrow Q_F < Q_G$. Since $1 - Pr(T = 1|\widehat{q} = 1) > 0$ we get that $Pr(T = 1|F) > Pr(T = 1|G)$.

The same applies to $I-$ideological agents. Now notice that if $|V(a) - V_I|$ is symmetric around zero, then we can compare I-ideological with O-ideological agents and derive that $|\pi_F| > |\pi_G| \Rightarrow Pr(T = 1|F) > Pr(T = 1|G)$.■

2.10 Appendix 2B: description of variables and regression results

2.10.1 Information derived from the British General Election Study 1997

• **INFO.**

The variable $INFO$ has been constructed by using the following two questions:

1. Do you happen to remember the names of any candidates who stood in your constituency in the general election this year?

Please write in all the names of candidates that you can remember (6 spaces provided) or tick box: I can't remember any of the candidates' names.

Note: the names of candidates written in by respondents were checked against official lists of candidates.

2. Political knowledge quiz (answers: true/false/don't know):

a: Margaret Thatcher was a Conservative Prime Minister;

b: The number of MP is about 100;

c: The longest time allowed between general elections is four years;

d: Britain's electoral system is based on proportional representation;

e: MPs from different parties are on parliamentary committees;

f: Britain has separate elections for the European parliament and the British parliament;

g: No-one may stand for parliament unless they pay a deposit.

Let us define with *names* the number of candidates correctly reported and with *quiz* the number of correct answers in question 2. $INFO$ is then given by

$$INFO = names + 0.66 \times quiz$$

The reason *quiz* has been downweighted is due to the fact that being true /false questions, it was possible for respondents to guess the answer without really knowing it,

while this is not possible for *names*. Therefore, using Bayes' rule we have

$$\Pr(know|correct) = \frac{\Pr(correct|know)}{\Pr(correct|know) + \Pr(correct|don't)} = \frac{1}{1+0.5} = 0.66$$

- **TNT.** (official turnout or declared turnout for those whose register was unavailable) 1 = voted.
- **income.** total household income from all sources before tax. Categorical variable from 1 to 16.
- **age.** respondent's age (>18).
- **age2.** = age^2×0.01.
- **sex.** 1 = male.
- **education.** respondent's education level. Categorical variable from 1 to 7.
- **married.** 1=yes (= 1 also if "living as married").
- **ethnicity**. "To which of these groups do you consider you belong?". **asian** = 1 if answer is one of "Indian, Pakistani, Bangladeshi, Chinese, Other Asian". **black** = 1 if answer is one of "Black African, Black Caribbean, Other Black".
- **churchgoer.** Categorical variable. "Apart from such special occasions as weddings, funerals and baptisms and son on, how often do you attend services or meetings connected with your religion?"

0. No religion or never or practically never attends;
1. varies too much to say;
2. less often than once a year;
3. at least once a year;
4. at least twice a year;
5. at least once a month;
6. at least once in two weeks;
7. once a week or more.

- **length of residence.** "How long have you lived in this neighbourhood?" (range 0-97).
- **farmer.** 1 if yes.
- **hours.** "How many hours (do/will/did) you normally work a week in your main job, including any paid or unpaid overtime?".

• **houseowner.** "Does your household own or rent this accommodation?". =1 if owns (leasehold etc.)

• **canvasser.** "Did a canvasser from any party call at your home to talk to you during the electoral campaign?". 1=yes.

• **phoned.** "Were you contacted by anyone on the telephone during the electoral campaign asking how you might vote?". 1=yes.

• **voted92.** =1 if voted in 1992 general election (self reported).

• **broadsheet-reader.** =1 if the answer is "yes" to

a: "do you regularly read one or more daily morning newspapers?"

and the answer to the question

b: "which daily morning newspaper do you read most often?" is one of the following:

- The Daily Telegraph;

- The Financial Times;

- The Guardian;

- The Independent;

- The Times.

• **economic activity.** Categorical variable:

1. "in paid work for at least 10 hours in week" or "waiting to take up paid work already accepted";

2. "in full time education (not paid for by employer, including on vacation)";

3. "on government training/employment programme";

4. "unemployed";

5."permanently sick or disabled";

6. "wholly retired from work";

7. "looking after the home";

8. "other"

• **union.**

Respondent or his/her partner is or has been member of a union. 1 if yes.

• **reg-i.**

General Standard Regions: i=1..11.

• **party effort in constituency.** Let us indicate with K the number of respondents in constituency j. For each respondent we know if she has been contacted by parties

(information in "canvasser"(c) and "phoned" (p)). Then for agent i in constituency j we have $c_{ij} \in \{0,1\}$ and $p_{ij} \in \{0,1\}$. We define party effort pe in constituency j as

$$pe_j = \frac{\sum_{i=1}^{K}(c_{ij} + p_{ij})}{2K} \in [0,1].$$

• **Left-Right1.** Derived from answers to the following question:

"In politics people sometimes talk of left and right. Where would you place yourself on a scale from 0 to 10, where 0 means the left and 10 means the right?"

Left-Right1 = 0 if answer is 5, Left-Right1 = 1 if answer is 4 or 6, Left-Right1 = 2 if answer is 3 or 7 etc. Respondents who answered "can't choose" are excluded.

• **Left-Right2.** Same as Left-Right1, but now respondents who answered "can't choose" are included with Left-Right2 = 0.

• **Party-identification.** Based on three questions.

1. "Do you generally think of yourself as a little closer to one of the parties than the others? If yes, which party?". Outcome: a) no; b) yes → [party named].

2. "Would you call yourself [party named] very strong, fairly strong or not very strong?

3. "Which one of the reasons on this card comes closest to the main reason you voted for the party you chose?"

Party = 1 if answer to question 1 is "no" or "don't know".

Party = 2 if answer to question 2 is "not very strong" or "don't know".

Party = 3 if answer to question 2 is "fairly strong".

Party = 4 if answer to question 2 is "very strong".

Party = 5 if answer to question 3 is "I always vote that way", independently of answers to questions 1 and 2.

2.10.2 Information about districts from Census 1991

• **higher education.** % of population with education qualification above university degree.

• **degree.** % of population with a degree but not higher education qualifications.

• **unemployed.** % unemployed.

• **population density.** Persons per hectare.

2.10.3 Information from Boothroyd (2002)

• **aggregate turnout**. Percentage turnout at the constituency level.

• **marginality.** Define with W and R the percentage of votes reported respectively by the winning candidate and the runner up. Then

$$\text{marginality} = \frac{W - R}{W + R}$$

2.10.4 Information from "The Guardian"

• **salience.** = 1 if an article specifically focused on a constituency electoral campaign appears on the Guardian between 1st and 30th April 1997.

2.10.5 Other

• **bbc100**. = 1 if constituency classified by the BBC among the 100 decisive constituencies (the battleground).

• **big shot**. = 1 if a current or former member of cabinet, a current member of shadow-cabinet or Paddy Ashdown is candidate in the constituency.

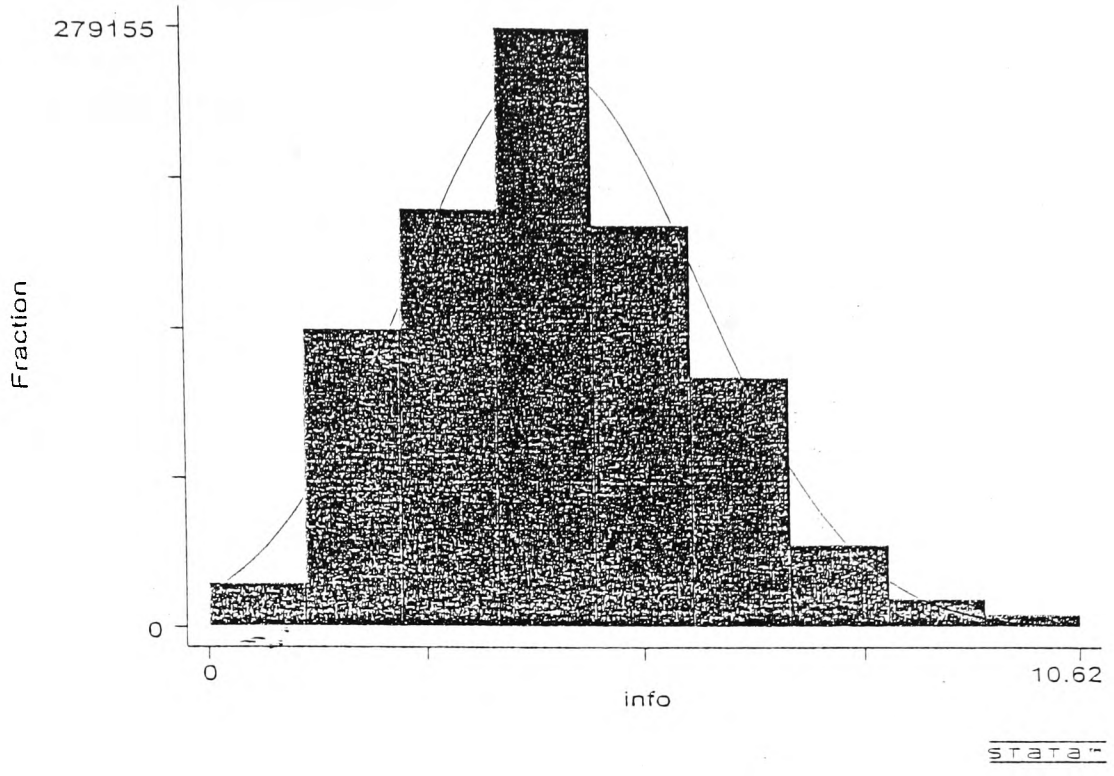


Figure 4-4: The distribution of *INFO*

Table 2.1: Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Turnout	2769	.794	.404	0	1
INFO	2769	4.306	1.8	0	10.62
age	2769	48.5	17.494	18	94
education	2769	3.597	2.161	1	7
income	2769	7.048	4.576	1	16
married	2769	0.59	0.492	0	1
sex	2769	0.466	0.499	0	1
asian	2769	0.018	0.132	0	1
black	2769	0.009	0.095	0	1
churchgoer	2769	1.99	2.61	0	7
union	2769	0.601	0.49	0	1
length of residence	2769	19.733	17.952	0	94
farmer	2769	0.006	0.076	0	1
hours	2769	38.219	15.837	0	95
house	2769	0.684	0.465	0	1
canvasser	2769	0.241	0.428	0	1
phoned	2769	0.075	0.263	0	1
voted92	2769	0.803	0.398	0	1
marginality	2769	0.303	0.194	0.005	0.814
aggregate turnout	2769	71.306	5.115	51.4	80
party effort in constituency	2769	0.156	0.124	0	1
Higher Education %	2769	.917	.791	.112	7.376
Degree %	2769	6.025	2.706	1.494	17.976
Unemployment ratio	2769	.093	.039	.029	.225
Density/1000	2769	133.909	166.972	0.219	1110.492
bbc100	2769	.133	.339	0	1
salience	2769	.094	.292	0	1
broadsheet-reader	2769	0.118	0.323	0	1
big-shot	2769	.068	.252	0	1
left-right1	2036	2.632	1.569	0	6
left-right2	2408	2.225	1.728	0	6
party-identification	2724	3.159	1.249	1	5

Table 2.2: Categorical Variables

Variable	Freq.	Percent	Cumulative
income			
less than 3999 £	217	7.84	7.84
4000-5999	347	12.53	20.37
6000-7999	244	8.81	29.18
8000-9999	190	6.86	36.04
10000-11999	214	7.73	43.77
12000-14999	238	8.6	52.37
15000-17999	194	7.01	59.37
18000-19999	137	4.95	64.32
20000-22999	177	6.39	70.71
23000-25999	163	5.89	76.6
26000-28999	129	4.66	81.26
29000-31999	95	3.43	84.69
32000-34999	79	2.85	87.54
35000-37999	52	1.88	89.42
38000-40999	65	2.35	91.77
41000 or more	228	8.23	100
education			
no qualification	937	33.84	33.84
foreign or other	16	0.58	34.42
CSE or equivalent	291	10.51	44.93
O level or equivalent	489	17.66	62.59
A level or equivalent	354	12.78	75.37
higher education below degree	381	13.76	89.13
degree	301	10.87	100
churchgoer			
no religion or never attends	1595	57.6	57.6
varies too much to say	33	1.19	58.79
less often than once a year	122	4.41	63.2
at least once a year	169	6.1	69.3
at least twice a year	291	10.51	79.81
at least once a month	140	5.06	84.87
at least once in two weeks	70	2.53	87.4
once a week or more	349	12.6	100

Continues on the next page

Table 2.2: Categorical Variables (continued)

economic activity			
paid work	1468	53.02	53.02
full time education	8	0.29	53.3
government training	64	2.31	55.62
unemployed	123	4.44	60.06
permanently sick or disabled	130	4.69	64.75
retired	640	23.11	87.87
looking after the home	318	11.48	99.35
doing something else	18	0.65	100
region			
North	151	5.45	5.45
North-West	202	7.3	12.75
Yorkshire & Humberside	206	7.44	20.19
West Midlands	241	8.7	28.89
East Midland	174	6.28	35.18
East Anglia	108	3.9	39.08
South West	193	6.97	46.05
South East	455	16.43	62.48
Greater London	228	8.23	70.71
Wales	132	4.77	75.48
Scotland	679	24.52	100
left-right1			
left-right1_1	672	33.01	33.01
left-right1_2	396	19.45	52.46
left-right1_3	410	20.14	72.59
left-right1_4	297	14.59	87.18
left-right1_5	89	4.37	91.55
left-right1_6	172	8.45	100
left-right2			
left-right2_1	1044	43.36	43.36
left-right2_2	396	16.45	59.8
left-right2_3	410	17.03	76.83
left-right2_4	297	12.33	89.16
left-right2_5	89	3.7	92.86
left-right2_6	172	7.14	100
party			
party1	162	5.95	5.95
party2	812	29.81	35.76
party3	826	30.32	66.08
party4	280	10.28	76.36
party5	644	23.64	100

Table 2.3: Turnout: logit coefficient estimates

(coefficients and marginal effects at the mean)

Dependent Variable: Turnout

	coeff	marg.eff.	z	coeff	marg. eff.	z	coeff	marg. eff.	z
information on individuals									
age	-.0038	-.0005	(0.16)	.0008	.00002	(0.03)	.0472	.0068	(2.03)
age2	.0071	.0010	(0.28)	.0019	.00039	(0.08)	-.0324	-.0047	(1.33)
education	.0693	.0098	(1.93)						
income	.0383	.0054	(1.96)						
married	.3590	.0530	(2.68)	.3034	.04597	(2.16)	.3665	.0554	(2.68)
sex	.0108	.0015	(0.08)	-.0094	-.00252	(0.07)	-.2009	-.0290	(1.55)
asian	.7808	.0857	(1.39)	.7114	.07002	(1.33)	.5001	.0617	(0.94)
black	.0206	.0029	(0.03)	-.1495	-.02559	(0.21)	.1120	-.0156	(0.15)
churchgoer	.0595	.0084	(2.28)						
union	.0950	.0135	(0.74)	.0567	.00711	(0.44)	.1170	.0170	(0.92)
length of residence	.0076	.0011	(1.89)	.0083	.00119	(2.04)	.0097	.0014	(2.40)
farmer	.9381	.0963	(1.18)	.9677	.09704	(1.22)	1.1164	.1106	(1.42)
hours	-.0125	-.0018	(2.72)	-.0130	-.00175	(2.82)			
houseowner	.2278	.0333	(1.61)	.2702	.04257	(1.78)	.3234	.0490	(2.29)
canvasser	.4117	.0545	(2.57)	.4361	.05583	(2.65)	.4264	.0576	(2.67)
phoned	.2780	.0363	(0.89)	.3145	.04256	(1.08)	.2733	.0366	(0.94)
voted 92	1.4564	.2614	(9.87)	1.4818	.26896	(9.92)			
constant	-2.8488		(1.72)	-3.5138		(2.05)	-4.6168		(2.68)
information on districts									
marginality	-.3521	-.0497	(0.78)	-.2822	-.03272	(0.64)	-.0310	-.0045	(0.07)
aggregate turnout	.0317	.0045	(1.69)	.0349	.00492	(1.81)	.0381	.0055	(2.11)
higher education %	-.0754	-.0106	(0.35)	-.0844	-.00935	(0.39)	.0068	-.0009	(0.04)
degree %	.0328	.0046	(0.55)	.0420	.00452	(0.70)	.0221	.0032	(0.41)
unemployed %	2.2872	.3231	(0.63)	2.7343	.29733	(0.74)	2.7899	.4030	(0.80)
population density	.0482	.0068	(0.64)	.0485	.00756	(0.65)	.0194	.0028	(0.26)
party effort	-.1521	-.0215	(0.24)	-.2152	-.03096	(0.34)	-.3628	-.0524	(0.61)
categorical variables (p-values of chi-test)									
education			No			0.1553			0.1493
income			No			0.0035			0.0137
churchgoer			No			0.0751			0.0104
economic activity			0.4862			0.4209			0.5855
region			0.9481			0.9196			0.9862
Observations:			2769			2769			2769
Log-L			-1221.68			-1193.00			-1268.01
Pseudo R2			0.1213			0.1419			0.0880

Note: z-satistics based on robust standard errors in parenthesis

Table 2.4: Turnout and information: logit coefficient estimates

(coefficients and marginal effects)

Dependent Variable: Turnout

	Logit			Logit			2-step Logit			2-step Logit		
	coeff	marg eff	z	coeff	marg eff	z	coeff	marg eff	z	coeff	marg eff	z
individual level												
info	.2719	.0371	(6.19)	.2696	.0356	(6.11)	.6557	.0865	*(2.01)			
1st stage fitted info										.6625	.0900	*(2.14)
1st stage residuals							-.3965	-.0523	*(1.12)			
age	-.0344	-.0047	(-1.37)	-.0305	-.0040	(1.22)	-.0752	-.0099	(2.03)	-.0763	-.0104	(2.04)
age2	.0307	.0042	(1.18)	.0266	.0035	(1.03)	.0625	.0082	(1.88)	.0637	.0087	(1.91)
education	.0083	.0011	(0.22)									
income	.0196	.0027	(0.95)									
married	.3750	.0537	(2.75)	.3139	.0432	(2.19)	.3123	.0429	(2.18)	.3022	.0427	(2.15)
sex	-.1683	-.0230	(-1.15)	-.1832	-.0242	(1.23)	-.4291	-.0566	(1.90)	-.4363	-.0593	(1.92)
asian	.8905	.0908	(1.53)	.8139	.0821	(1.49)	1.1206	.1019	(1.80)	1.2240	.1113	(1.97)
black	.0042	.0006	(0.01)	-.1241	-.0171	(0.17)	-.1074	-.0147	(0.15)	-.1210	-.0171	(0.17)
churchgoer	.0524	.0072	(1.98)									
union	.0518	.0071	(0.40)	.0212	.0028	(0.16)	-.0559	-.0073	(0.39)	-.0722	-.0098	(0.50)
length of residence	.0066	.0009	(1.60)	.0070	.0009	(1.68)	.0045	.0006	(1.00)	.0041	.0005	(0.92)
farmer	.8974	.0899	(1.13)	.8541	.0838	(1.10)	.8045	.0801	(1.01)	.8796	.0881	(1.07)
hours	-.0114	-.0016	(2.48)	-.0118	-.0016	(2.55)	-.0096	-.0013	(1.98)	-.0093	-.0012	(1.93)
houseowner	.1955	.0275	(1.35)	.2247	.0308	(1.47)	.1543	.0208	(0.94)	.1507	.0210	(0.93)
canvasser	.4185	.0534	(2.58)	.4404	.0542	(2.64)	.4193	.0516	(2.48)	.3976	.0506	(2.38)
phoned	.2569	.0326	(0.84)	.2787	.0339	(0.97)	.1855	.0232	(0.64)	.1609	.0209	(0.54)
voted 92	1.3673	.2365	(9.13)	1.4001	.2376	(9.25)	1.2253	.2021	(6.18)	1.1836	.1986	(6.05)
district level												
marginality	-.1307	-.0178	(0.28)	-.0521	-.0069	(0.12)	.2663	.0351	(0.53)	.2626	.0357	(0.53)
aggregate turnout	.0308	.0042	(1.63)	.0343	.0045	(1.79)	.0302	.0040	(1.56)	.0279	.0040	(1.45)
higher education %	-.1150	-.0157	(0.50)	-.1254	-.0166	(0.55)	-.1939	-.0256	(0.86)	-.1980	-.0269	(0.91)
degree %	.0374	.0051	(0.60)	.0472	.0062	(0.76)	.0615	.0081	(0.99)	.0658	.0089	(1.08)
unemployed %	.7857	.1073	(0.21)	1.2695	.1679	(0.34)	-0.7076	-.0933	(0.18)	-.7152	-.0972	(0.18)
population density	.0744	.0102	(0.99)	.0789	.0104	(1.06)	.1140	.0151	(1.45)	.1110	.0150	(1.42)
party effort	-.3075	-.0419	(-0.49)	-.3373	-.0446	(0.53)	-.5418	-.0715	(0.81)	-.5782	-.0786	(0.86)
constant	-5.59355		(-3.16)	-3.0322		(1.77)	-2.1940		(1.23)	-2.0628		(1.16)
categorical variables (p-values of chi-test)												
education		No			0.6553			0.5893			0.6080	
income		No			0.0119			0.0195			0.0209	
churchgoer		No			0.1235			0.2289			0.2297	
economic activity		0.8190			0.7576			0.8728			0.8068	
region		0.8241			0.7669			0.5776			0.6336	
Observations:		2769			2769			2769			2769	
Log-L		-1192.74			-1165.70			-1164.28			-1188.78	
Pseudo R2		0.1421			0.1616			0.1626			0.1450	

Note: z-statistics from robust standard errors in parenthesis. Z-statistics marked with * are calculated by bootstrap

Table 2.5: First stage regression: OLS coefficients
(t-statistics in parentheses)

Dependent Variable: Information

	coeff	t
age	.1121	(8.16)
age2	-.0907	(6.53)
married	.0069	(0.08)
sex	.6155	(7.44)
asian	-.7722	(3.65)
black	-.0170	(0.05)
union	.2380	(3.32)
length of residence	.0072	(3.22)
farmer	.1027	(0.26)
hours	-.0051	(-1.91)
houseowner	.1914	(2.26)
canvasser	.0902	(1.07)
phoned	.1839	(1.23)
voted 92	.4578	(4.89)
marginality	-.7047	(2.84)
aggregate turnout	.0064	(0.65)
highed education %	.1429	(1.77)
degree %	-.0339	(1.20)
unemployed %	4.2717	(2.13)
population density	-.0774	(1.79)
Party Effort	.2371	(0.67)
constant	-1.7663	(2.00)
Instrumental variables		
bbc100	.2945	(2.55)
salience	.1853	(1.60)
quality-paper reader	.7163	(6.79)
big shot	.2425	(1.62)
categorical variables (p-values of F-test)		
education		0
income		0.0464
churchgoer		0.5801
economic activity		0.0005
region		0.0019
Observations:		2769
R2		0.3161

Note: Robust standard errors

Table 2.6: Testing the over-identification restrictions

(Likelihood-ratio test)

Dependent Variable: Turnout

	coeff	z	coeff	z
info			.26024	(5.77)
bbc100	.37268	(1.64)	.31473	(1.38)
salience	.13742	(0.62)	.07651	(0.34)
quality-paper reader	.39284	(1.68)	.1843	(0.78)
big-shot	.19252	(0.73)	.17411	(0.64)
other control variables	Yes		Yes	
Observations:		2769		2769
Log-L		-1188.28		-1163.61
Pseudo R2		0.1453		0.1631
L-Ratio statistics		1.00		4.18

Table 2.7: Information and Ideology: OLS coefficients
(t-statistics in parentheses)

Dependent Variable: Information

	coeff	t	coeff	t	coeff	t
age	.0776	(4.73)	.0888	(6.01)	.1141	(8.30)
age2	-.0590	(3.51)	-.0679	(4.51)	-.0930	(6.68)
married	-.0071	(0.08)	.0076	(0.09)	.0077	(0.09)
sex	.5103	(5.51)	.5666	(6.48)	.6170	(7.44)
asian	-.7933	(2.86)	-.7556	(3.20)	-.7904	(3.69)
black	.4382	(1.09)	-.3342	(0.85)	.0110	(0.03)
union	.3037	(3.78)	.2513	(3.37)	.2099	(2.92)
length of residence	.0090	(3.59)	.0096	(4.06)	.0076	(3.40)
farmer	-.0369	(0.10)	.0161	(0.04)	.0754	(0.19)
hours	-.0054	(1.79)	-.0049	(1.75)	-.0049	(1.82)
houseowner	.3049	(3.04)	.2000	(2.20)	.1883	(2.22)
canvasser	.1283	(1.38)	.1814	(2.07)	.0987	(1.17)
phoned	.2346	(1.43)	.2657	(1.78)	.1730	(1.14)
voted 92	.3342	(3.13)	.3470	(3.62)	.4503	(4.46)
marginality	-.6349	(2.25)	-.7579	(2.89)	-.6752	(2.71)
aggregate turnout	.0061	(0.57)	.0090	(0.90)	.0035	(0.35)
highed education %	.0773	(0.93)	.1110	(1.40)	.1390	(1.70)
degree %	-.0150	(0.49)	-.0290	(1.02)	-.0293	(1.04)
unemployed %	2.7366	(1.21)	3.8615	(1.87)	4.0696	(2.05)
population density	-.0538	(1.08)	-.0648	(1.41)	-.0838	(1.96)
party effort	.3064	(0.74)	.3771	(0.99)	.2798	(0.79)
bbc100	.2582	(2.01)	.2057	(1.70)	.3076	(2.64)
salience	.1999	(1.60)	.1789	(1.55)	.1620	(1.40)
quality-paper reader	.6434	(5.78)	.6743	(6.17)	.7224	(6.80)
big shot	.3557	(2.16)	.3110	(2.03)	.2201	(1.47)
*left-rightx_2	.1455	(1.40)	.4031	(4.13)		
*left-rightx_3	.2297	(2.21)	.5005	(5.10)		
*left-rightx_4	.1236	(1.04)	.3899	(3.42)		
*left-rightx_5	.0203	(0.10)	.3200	(1.62)		
*left-rightx_6	-.1401	(0.89)	.1451	(0.94)		
party attachment 2					.0788	(0.55)
party attachment 3					.2336	(1.54)
party attachment 4					.4095	(2.28)
party attachment 5					.0443	(0.28)
constant	-.3197	(0.32)	-1.1936	(-1.30)	-1.7029	(1.93)
categorical variables (p-values of F-test)						
education		0		0		0
income		0.0577		0.1089		0.0612
churchgoer		0.2433		0.4622		0.5037
economic activity		0.1318		0.1496		0.0012
region		0.0001		0.0006		0.0011
Observations:		2036		2408		2724
R2		0.2963		0.3250		0.3239

Note: Robust standard errors

*Left-Right1 reported in column 1 and Left-Right2 reported in column 2

Table 2.8: Turnout, ideology, and information: logit estimates

(coefficients and marginal effects)

Dependent Variable: Turnout

	coeff	marg eff	z	coeff	marg eff	z	coeff	marg eff	z
individual level									
info	.3704	.0426	(3.79)	.29430	.03666	(4.41)	.60914	.07547	(5.89)
age	.0166	.0019	(0.54)	.01030	.00128	(0.38)	-.00742	-.00092	(0.29)
age2	-.0262	.0030	(0.84)	-.01690	-.00211	(0.60)	.00045	.00006	(0.02)
married	.2588	.0310	(1.49)	.29494	.03834	(1.90)	.32167	.04163	(2.15)
sex	.0257	.0030	(0.14)	-.04516	-.00562	(0.28)	-.07790	-.00965	(0.50)
asian	-.1625	-.0198	(0.27)	.43806	.04691	(0.67)	.64684	.06430	(1.16)
black	-.9725	-.1545	(1.05)	-.36386	-.05128	(0.40)	-.47081	-.06832	(0.54)
union	-.0533	-.0061	(0.33)	-.03917	-.00486	(0.27)	-.05890	-.00727	(0.43)
length of residence	.0066	.0008	(1.30)	.00796	.00099	(1.75)	.00356	.00044	(0 81)
farmer	1.9989	.1135	(2.36)	1.86067	.12103	(2.06)	.88407	.07979	(1.10)
hours	-.0150	-.0017	(2.58)	-.01480	-.00184	(2.93)	-.01371	-.00170	(2.76)
houseowner	.2115	.0253	(1.09)	.19304	.02483	(1.15)	.32596	.04259	(2.02)
canvasser	.5304	.0555	(2.52)	.41328	.04791	(2.25)	.36096	.04206	(2.07)
phoned	.2416	.0257	(0.71)	.30110	.03424	(0.96)	.25470	.02921	(0.84)
voted 92	1.0821	.1596	(5.39)	1.23010	.19642	(7.21)	1.0232	.15562	(6.18)
*ideology	.3484	.0400	(2.09)	.21102	.02629	(1.77)	1.06171	.13155	(7.19)
*infoXideology	-.0440	-.0051	(1.15)	-.02368	-.00295	(0.80)	-.12880	-.01596	(3.68)
district level									
marginality	-.1859	-.0214	(0.33)	-.26154	-.03258	(0.52)	-.15180	-.01881	(0.32)
aggregate turnout	.0355	.0041	(1.50)	.02502	.00312	(1.15)	.02499	.00310	(1.26)
higher education %	-.2672	-.0307	(1.14)	-.17599	-.02192	(0.72)	-.13420	-.01663	(0.58)
degree %	.0824	.0095	(1.18)	.06800	.00847	(1.00)	.04454	.00552	(0.69)
unemployed (ratio)	1.9007	.2186	(0.41)	1.48528	.18502	(0.35)	-.82380	.10207	(0.21)
population density	.0548	.0630	(0.59)	.04230	.05270	(0.52)	.06280	.07780	(0.80)
party effort	-.0043	-.0005	(0.01)	-.25340	-.03157	(0.36)	-.25543	-.03165	(0.38)
constant	-5.0932		(2.35)	-3.52188		(1.80)	-5.20368		(2.90)
categorical variables (p-values of chi-test)									
education		0.9004			0.6333			0.4999	
income		0.077			0.061			0.0115	
churchgoer		0.0570			0.1735			0.1424	
economic activity		0.7973			0.7302			0.8441	
region		0.3464			0.2691			0.6208	
Observations:		2036			2408			2724	
Log-L		-798.87			-980.022			-1073.98	
Pseudo R2		0.1558			0.1652			0.2141	

Note: z-statistics from robust standard errors in parenthesis.

* Ideology is Left-Right1 in column 1, Left-Right2 in column 2, and Party Identification in column 3.

Chapter 3

Rational Ignorance and the Public Choice of Redistribution

3.1 Introduction

Since the early stages of the economic theory of politics, Downs pointed out that in a sizeable electorate "the returns from voting are usually so low that even small costs may cause many voters to abstain". This carries implications not only for political participation but also for the desire to be informed about political issues. If there is a cost of acquiring information about the candidates and their platforms, then we should expect not only rational abstention but also "rational ignorance" on political issues.

This consideration implies a substantial lack of information by citizens about candidates and their proposals. The fact that many people actually vote and that political information is still available in newspapers would be simply reduced to a matter of preferences: political information may be enjoyable *per se*, not unlike sports news[1]. This is equivalent to admitting that preferences for political information, like all preferences, are outside the domain of standard economic theory[2]. If this was true then the chances of

[1] Analogously, Riker and Ordershook (1968) explain voters' turnout in general elections by including a sense of citizen's duty in individuals' preferences.

[2] In the words of Downs, "a rational man can become well informed for four reasons: 1) he may enjoy being well informed for its own sake, so that information as such provides him with utility; 2) he may believe the election is going to be so close that the probability of his casting the decisive vote is relatively high; 3) he may need information to influence the votes of others (...); 4) he may need information to influence the formation of government policy as a lobbyist. Nevertheless, since the odds are that no election will be close enough to render decisive the vote of any one person, or the votes of all those he can persuade to agree with him, the rational course of action for most citizens is to remain politically

being informed or of showing up at the voting booth could be expected to be independent of most economic variables, which seems to be at odds with most empirical research[3].

From the previous chapter it should be clear that, apart from the obvious role of personal preferences, the demand for political information can be explained in terms of incentives. Here we will derive the consequences of this idea for redistributive policy-making.

We will argue that rational ignorance is the consequence of an artificial separation between politics and the economy. It seems instead intuitive that expectations on policies should be relevant to private decisions. This generates a demand for political information to be used for private purposes. Under quite mild assumptions, this demand is positively correlated with income: in other terms, we can expect the rich to be systematically better informed than the poor, independently of any demand for information purely as a consumption good.

From the study of the U.K. 1997 election of the previous chapter, we know already that political knowledge is increasing with income. Other research points clearly, and not surprisingly, in the same direction[4]. This observation adds an important element in the political market and can help to explain the poor empirical support encountered by Downsian theories of redistribution.

According to voting models of redistribution based on the median voter theorem, income inequality should increase redistribution as long as it increases the distance between average income and the income of the pivotal voter (Roberts,1977); this result has been applied to a variety of situations to explain the size of the public sector, low growth rates, increasing intergenerational transfers and so on. However, it is also fair to say that this theory does not receive good support from empirical research. Even though the reduced forms referring to specific situations are generally compatible with the data, when moving to structural-form analysis (linking inequality to some measure of redistributive transfers), support is generally weak and coefficients often show signs different from those expected[5]. There are various possible explanations for this unsatisfactory empirical sup-

uninformed" (Downs, 1957).

[3]See for example Matsusaka (1995) and the references given there.

[4]See Delli Carpini and Keeter (1996).

[5]For examples of reduced form analysis see Alesina and Rodrick (1994) and Persson and Tabellini (1994). Estimations of structural relationships between redistributive transfers and inequality are given in Perotti (1994) and Lindert (1996).

port[6]; however, it seems clear that the theory, though representing a useful benchmark, provides a simplistic representation of how democratic systems work. Other institutional elements and country-specific features are likely to affect the policy outcomes.

It is worth remembering that this benchmark depends on some crucial assumptions that have been challenged on a variety of grounds. First of all, it requires unidimensionality of the policy space. When public policy is considered in a multidimensional space, then an equilibrium may not exist or it may assume very different characteristics[7]. In particular, political platforms proposed by candidates do not necessarily converge. Also, the median voter theorem requires that political parties be perfectly able to commit to their proposed policies. When candidates are unable to make credible commitments then the tendency to platform divergence in equilibrium is reinforced[8].

This study points in a different direction. The model unveils a possible relationship between incentives to gather political information and preferences over redistribution. Information acquisition might be non-neutral for voting outcomes: indeed, our model implies a substantial heterogeneity in awareness on policies, which could affect political competition and eventually policy choices. It will be shown how this may provide a possible explanation of the weak empirical support for the traditional benchmark.

We focus on the demand for political information. It is clear that in an economic theory of politics there is no simple explanation for any type of political participation, where participation must be taken in the broad sense of voting, taking part in political organizations, acquiring political information and so on.

One first possibility, as noted, is that political information is demanded as a consumption good and not for decision-making: most people seem to enjoy being informed on many things, even when this does not enable them to make better decisions. In this case one should ask about the nature of this good and, in particular, whether it is a normal good. This is clearly an empirical matter; if, as we have seen, political information can be treated as a normal good, then the rich can be expected to be more informed than the poor and therefore more responsive to policy announcements: all the results we present in this work would be valid *a fortiori*.

[6] Among other things, it is worth remembering that for some countries data are not completely reliable.

[7] See for example Besley and Coate (1997).

[8] See Alesina (1987 and 1988) for partisan models of two-party electoral competition. Besley and Coate (1997) also consider policy-oriented citizen-candidates.

In this chapter, however, we refer only to information as it is considered in decision theory, ignoring information as a consumption good and not relying on normality. In fact, the premise that political information is rarely relevant to useful decision-making relies on an artificial modelling separation between politics and the economy. Our working assumption, instead, is that political information may be acquired for private purposes and that this incentive is relevant.

Many pieces of information may be relevant when voting even though they were acquired for some other purpose. For example, information on fiscal variables may be relevant to investment decisions and at the same time convey information on economic policy; information on the quality of some public service (for example health) may be useful to know whether it is worthwhile using privately available alternatives and at the same time can reveal information on the effort of the current administration to provide good services. Moreover, at election time, political information may be acquired to form more accurate expectations on future policy: investment decisions today depend on expectations on future taxes; choosing a public or a private school today involves expectations over the condition of the educational system in a few years; and so on.

Sometimes the behaviour of political agents may reveal, apart from policies, something about the external world that is relevant to private decision making. Political parties have every incentive to collect information for their own action, so accurate observation of their choices can convey information on many variables that are unobservable (or too costly to observe) to the private citizen.

The analysis presented in the following builds on the theory of information acquisition presented in the previous chapter, although the role of ideology will be neglected. Information therefore will have some characteristics not often considered in the literature. First of all, information does not come effortlessly: agents must spend effort and time to gather and process information. Secondly, acquiring information is an activity with uncertain returns: more time and effort makes it more likely to get better information, but there is no certainty about what and how much is going to be known. Third, information is considered as freely accessible to all: this makes our analysis particularly suited for information available in the mass media. In fact, the revenue of most newspapers and broadcasts comes from advertising: attracting a larger public raises the value of units to

sell to advertisers[9]. The consumer in this case does not pay information in cash; in any event, this cost is quite low compared with some other opportunity cost.

It is important to stress that defining the value of information and deriving a demand for it requires dealing with some problems posed by its special characteristics. First, information demand is a derived demand: information is valuable because it enables people to make better choices[10]. This means that information cannot be put in the commodity space when defining preferences. As a consequence, relevant nonconcavities may arise to complicate the analysis, leading to an unsatisfactory theory overall[11]. Second, to specify a model of information demand we need a clear definition of the information available, its costs and the decision making process. Information is valuable only when there is uncertainty on variables that are relevant to decision making. Third, there is no easy way to define the quantity of information. Given a space of possible states Ξ, we can say that signal ξ is more informative than signal ξ' when it induces a finer partition of the state space: but this does not provide a complete order of signals, as many partitions are simply not comparable with this criterion. Thus, a complete ordering of signals may be obtained only with reference to a score function, i.e. with reference to how the signals are valuable in terms of the decisions to be made: this means that there is no objective, permanently valid definition of the quantity of information in economics[12].

Most voting models with asymmetric information have typically considered either a representative voter imperfectly informed on candidates (e.g. Harrington 1993) or fixed political alternatives (e.g. Palfrey and Rosenthal, 1985, Feddersen and Pesendorfer, 1996). Ledyard (1984) presents a model of spatial electoral competition where each voter is uncertain about other voters preferences and cost of voting, and where abstention is admitted. Voters play a Bayesian game for given candidates' positions; this gives positive turnout when candidates' positions are differentiated. Candidates, however, are lead to convergence by competition for votes and this drives the equilibrium turnout to zero. In McKelvey and Ordeshook (1984) some voters are uninformed about the candidates' positions but they know the preferences of the various subgroups in the population; un-

[9] See the discussion of this point in Stromberg (2002).

[10] We are referring to the notion of information in decision theory. All other information can clearly be included in the category of leisure.

[11] See for example Radner and Stiglitz (1984).

[12] The Shannon measure of the quantity of information, derived in a different context, has proved to be of little use in economic theory. See Shannon (1948).

informed voters can make inferences using interest-group endorsement and opinion polls. Under certain assumptions about preferences and preference distribution, all voters choose as if they had perfect information. McKelvey and Ordeshook conclude that perfect information is not a necessary condition to apply the median voter theorem. Stromberg (2002) introduces mass media as information sources: since some voters are more valuable than others to advertisers they will get better coverage of the issues they are interested in. Electoral competition between office-seeking candidates will then translate the mass media bias into a policy bias.

There is no model, to my knowledge, that introduces the idea of increasing returns to information into the political market. This idea is clearly not new in other applications. Among those, Arrow (1986) is of particular relevance for our analysis. In Arrow, information is demanded for portfolio decisions under uncertainty. The analysis is limited to this specific case and considers a given specification for the utility function (CES). Information is provided by a signal on returns, and the quality of the signal is given by its precision. Arrow concludes that different incentives to acquire information (the asymmetry between fixed costs and increasing returns) lead portfolio allocation choices to increase income inequality. Other studies on the demand for information include Kihlstrom (1974) and Verrecchia (1982). Kihlstrom (1974) provides a general theory of information demand about product quality, when consumers are interested not directly in commodities but in some desirable attributes they may have. The quantity of information is defined using, as in Blackwell (1951), the concept of sufficiency: if an observable random variable ξ is sufficient for $\xi^{'}$, then ξ delivers more information than $\xi^{'}$. In Verrecchia (1982) agents may acquire private signals about stocks' returns on top of what equilibrium prices already reveal, but there is no wealth effect. In these models the cost of a better signal is a monetary cost.

The model of information demand of this chapter will be rather simple, neglecting many of the complications of chapter 2. This is necessary to keep the analysis manageable as we will now develop a parallel model of electoral competition and policy formation. However, this simplified framework should be enough to show the relevance of political information for public choice and public policy.

The chapter is organized as follows. The next section presents a simple model of Downsian political competition in which both private and public decisions must be made

by citizens. In Section 3.3 we derive the demand for political information and show that incentives to be informed on politics are increasing in agents' initial endowments. In Section 3.4 we solve the model and analyse the role of information on political equilibrium. Section 3.5 discusses the main implications of the model for the interaction between gross income inequality and redistribution. Section 3.6 briefly discusses the main normative issues at stake in this analysis and the role of coordination failures in information acquisition. Section 3.7 concludes.

3.2 The framework of the model

In the following model political competition is limited to a Downsian two-party system with full commitment. Of course this implies that the model has all the limitations of the Downsian analysis, which we do not intend to focus on here. It is instead important to compare our results with a standard Downsian model. Even though the analysis is kept as simple as possible, this does not preclude the applicability of this framework to more sophisticated models of political competition.

Our economy consists of a continuum of agents. Each agent's preferences will be represented by a continuous utility function

$$u(\mathbf{x}, e, a|m) = U(\mathbf{x}|a) + Z(a|m) - ve \tag{3.1}$$

where $\mathbf{x}$ is a vector of private goods (with prices $\mathbf{p}$), $a \in A \equiv [\underline{a}, \overline{a}]$ is a public policy variable (e.g. a public good), m is the initial endowment, and $e \in \mathcal{E}$ is effort devoted to information gathering, with $v = \omega + \varepsilon$ a parameter of effort disutility. We assume ω to be a cost that is common to the whole population and distributed according to the function $p_\omega(\varpi, \sigma_\omega^2)$ with $S_\omega = \left\{\omega | p_\omega(\varpi, \sigma_\omega^2) > 0\right\} \subset \Re_+$; $\varepsilon \sim p_\varepsilon(\overline{\varepsilon}, \sigma_\varepsilon^2)$ is an idiosyncratic shock with $S_\varepsilon = \left\{\varepsilon | p_\varepsilon(\overline{\varepsilon}, \sigma_\varepsilon^2) > 0\right\} \subset \Re_+$. The function $U(\mathbf{x}|a)$ is the utility associated with private commodities. This is assumed to be contingent on the value of the public policy variable: for example the utility of buying a car depends on the quality of roads. Thus, knowing the value of a is important for the choice of the bundle $\mathbf{x}$. However, we also assume that people have direct preferences over a represented by a strictly concave function $Z(a|m)$. Since we want to focus on redistributive politics, we will assume that

preferences on a depend on agents' initial endowment.

We assume people have an identical utility function over private commodities $U(\mathbf{x}|a)$: hence the only ex ante source of heterogeneity is their initial endowment. An agent with endowment m has a choice set given by

$$\mathcal{X}_m = \{\mathbf{x}|\mathbf{px} \leq m(1 - \pi e)\} \tag{3.2}$$

where π is a positive parameter, equal for all agents, reflecting the possible monetary costs induced by information gathering (for example, via a reduction in labour supply). Since the maximum amount that can be spended in information gathering is m we have $e \in \mathcal{E} = [0, \frac{1}{\pi}]$. Interpreting the initial endowment as full income, we will summarize income distribution in the population by a continuous density function $\varphi(m)$.

Let us focus on the first component of the utility function, neglecting for the moment both $Z(a|m)$ and the choice of e. Let us also assume that a is fixed and known with certainty. We make the following assumption:

Assumption 3.1 *$U(\cdot) \in \Re_+$ is quasi-concave and homogeneous of degree 1 in $\mathbf{x}$.*

Although Assumption 3.1 clearly restricts the behaviour pattern of our agents, it should be noted that the class of utility functions we consider is still fairly general, comprising some of the standard functions most widely used in economic models.

From the constrained maximization of the utility function we get the optimal private choice $\mathbf{x}^*(a, m, \mathbf{p})$ and the indirect utility function $V(a, m, \mathbf{p})$. Preferences over a are then defined by the function $W(a, m, p) = V(a, m, \mathbf{p}) + Z(a|m)$. Each agent has therefore an ideal level of a defined by the function $W(a, m, p)$ and, for given prices, this depends on m. We will also assume the following:

Assumption 3.2 *$W(a, m, p)$ satisfies the single crossing condition: $\forall a' > a$, $\forall m' > m$: $W(a', m', p) \geq W(a, m', p) \Rightarrow W(a', m, p) \geq W(a, m, p)$ and $W(a', m', p) > W(a, m', p) \Rightarrow W(a', m, p) > W(a, m, p)$.*

Assumption 3.2 implies that richer agents prefer lower levels of a than the poorer ones. Given the continuity of the functions involved, we can represent the preferred policy of an agent with income m as a function $a = z(m)$ with $z' < 0$. We can think of a as any

policy issue; we only require preferences on a to be somehow related to income; thus, a could be some specific type of public good or a redistributive transfer in a second-best environment[13]. For the rest of this section we indicate the distribution of the ideal a (the argmax of the function $W(a, m, p)$) across the population with $y(a)$.

So far, our agents act on the economy by their private decisions; however, they may also affect the public decision with their votes. From now on we also assume that a is unknown. In our environment there are two parties (L and R) competing for office. They are able to commit to their platforms and care only about maximizing votes. Thus they have no preference for any platforms: these are used only instrumentally to convince voters.

Parties' platforms are announced publicly but are observable only if some effort e is devoted to information gathering. More precisely, we will assume that the probability of observing the vector of announcements $\{a_L, a_R\}$ is given by $q(e)$, where $q(\cdot)$ is an increasing and concave function. One possible interpretation of this assumption is that parties' communications are very often transmitted to voters only indirectly, by the mass media. Also, political platforms are very complex and the ultimate effect on an agent's finances is never very clear. Researchers use quite sophisticated models to approximate the effects of simple policies, so there is no reason why a voter should completely and immediately understand political platforms and their consequences. Although this critique could be extended to many other models in economics, it seems particularly relevant when we come to public policies, because of their intrinsic complexity.

The timing of the model is represented in figure 3-1: first of all Nature selects ω for the whole community and the idiosyncratic shocks ε for each citizen. Citizens only learn their own v. Politicians, however, may observe the realization ω. Both citizens and politicians know the distribution of policy preferences. In period 1 the two parties simultaneously announce their platforms. Citizens spend their desired amount of effort in acquiring information and afterwards decisions are made, i.e. private choices are undertaken and

[13] One possible situation leading to this framework is the choice of the tax rate in a proportional tax system with lump sum transfer and balanced budget. This is the situation analysed in Roberts (1977) and Meltzer and Richard (1981). It should be noted that in this case the desired level of a will depend on the ratio $\frac{m}{\tilde{m}}$, where $\tilde{m}$ is mean income. In other terms, the desirability of redistribution for an agent with income m is decreasing in each agent's own income and increasing in mean income. The policy preferred by the pivotal voter (usually with income below the mean) will depend on the distance of that voter's income from the mean. Therefore, in our electoral model m can always be replaced by $\frac{m}{\tilde{m}}$. For us this has no consequence since we will only compare distributions with the same mean income.

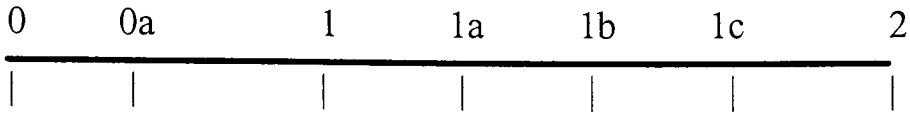


Figure 3-1: Time Line

0: Nature selects realization of ω for the whole population and ε for each citizen.

0a: Political parties learn ω, citizens learn v.

1: Parties simultaneously and independently announce political platforms.

1a: Citizens gather information on platforms.

1b: Private decisions.

1c: Voting decisions.

2: Winning platform is implemented. Payoffs realized.

people cast their votes on the basis of the information they have. Finally the announced policy of the winner party is implemented and payoffs are realized for all citizens.

Note that the model can easily accommodate a series of complications that would not change anything substantial. First of all, other sources of uncertainty could be added with no significant consequences. For example preference distribution $y(a)$ could be uncertain. If there are two possible distributions $y_1(a)$ and $y_2(a)$ with respective probabilities p and $(1-p)$ then a state of the world would be defined by realizations of information costs and preference distribution. An agent could learn something by observing his or her own preferences but would still be substantially uncertain, making information valuable. This possibility will be considered in example 2.

Another possibility is to allow only for the observation of a signal ξ on platforms, rather than the platforms themselves. In this case, assuming that the joint distribution of a and ξ satisfies the monotone likelihood ratio property, knowing ξ would reduce uncertainty and the set of possible political equilibria, still making information gathering an activity with positive returns. Note also that for our purposes the following analysis

would be the same if the function $U(\cdot)$ was represented as $U(\mathbf{x}|\psi(a))$ where $\psi(a)$ is any variable relevant to private decision-making and affected by public policies (for example, the interest rate)[14].

We will now start with the presentation of the information demand given its central role in this model. Then we will proceed to solve the model backward.

3.3 Private decisions and the demand for information

Private decisions are of two types: the choice of a commodity bundle $\mathbf{x}$ and the choice of e. As will become clear, these two choices must be analysed separately, as the choice of e requires defining a notion of the value of information and this, in turn, can be defined only with respect to the maximum value function, when private choices have been made. Therefore, a two-step maximization process will be used. Solving the individual decision-making process backward, we start by considering e fixed and equal to $\widetilde{e}$. Then we can temporarily ignore the role of e and π.

As we said, the public policy variable a is relevant to private decision making. Since the decision has to be made before (or simultaneously to) the election, a is unknown. The motivation for information gathering is to make better private decisions. However, since private choices depend on policies, it is convenient to start with political decisions.

A platform announcement by parties L and R is defined as a pair $\{a_L, a_R\}$. Every announcement will induce a partition of the whole population: let us indicate with $N_L(a_L, a_R)$ and $N_R(a_L, a_R)$ the size of the population that, if informed on the content of platforms, would vote respectively for party L and party R when $\{a_L, a_R\}$ is received.

However, not all the people in $N_L(a_L, a_R)$ and $N_R(a_L, a_R)$ will be informed on the platforms. Since there are no priors on parties' location, L and R are just labels, and therefore uninformed citizens are not responsive to parties' proposals; we will interpret this non-responsiveness as abstention, by assuming that any indifferent voters simply do not vote. Actually, in our setting there is not much an uninformed voter can do apart from voting randomly or abstaining. We then indicate with $n_L(a_L, a_R)$ and $n_R(a_L, a_R)$ the size of the informed population voting for party L and party R respectively when $\{a_L, a_R\}$ is received, and with $En_L(a_L, a_R)$ and $En_R(a_L, a_R)$ their respective expected

[14] This is straightforward if $\psi(a)$ is a monotonic function.

values when the size of the informed population is uncertain.

Let us indicate with $P^W(a_i|a_i,a_j)$ the probability that the platform of party i wins given that the platforms announced are $\{a_i,a_j\}$. Then we have

$$P^W(a_i|a_i,a_j) = \left\{ \begin{array}{l} 1 \text{ if } n_i(a_i,a_j) > n_j(a_i,a_j) \\ \frac{1}{2} \text{ if } n_i(a_i,a_j) = n_j(a_i,a_j) \end{array} \right\} \tag{3.3}$$

Information is used by our agents in the best possible way; we also assume that each citizen knows the distribution of public policy preferences $y(a)$. Therefore agents are able to infer the population partitions induced by any platform announcements. Since information is acquired to forecast future policies we have the following assumption about the expected policy:

Assumption 3.3

$$a^* = E(a|a_i,a_j) = \left\{ \begin{array}{c} a_i \text{ if } En_i(a_i,a_j) > En_j(a_i,a_j),\ i,j = L,R \\ \frac{1}{2}a_i + \frac{1}{2}a_j \text{ if } En_i(a_i,a_j) = En_j(a_i,a_j) \end{array} \right\}$$

In other words, each agent knows the population partitions induced by any pair of platform announcements, and, if informed about the platforms, can then forecast the future policy. This means that the optimal private decision $\mathbf{x}^*$ can be made contingent on $\{a_L,a_R\}$.

We will show later that the winning platform will depend on the realization of ω. For the moment let us just assume that the winning platform can be represented as a continuous function $a^*(\omega)$ (this will be proved in Lemma 3.1).

Focusing for the moment only on the choice of commodities (i.e. on the first component of the utility function), we have that the utility of an agent who observes the platform announcements is[15]

$$U(\mathbf{x}^*(m,a^*(\omega))|a^*(\omega)) \tag{3.4}$$

whereas if platforms have not been observed utility is

$$U(\widetilde{\mathbf{x}}(m)|a^*(\omega)). \tag{3.5}$$

Note also that when v is learned by each agent at the beginning of the game, the prior

[15] From now on we drop prices, as they do not vary in our analysis.

probability of ω can be updated by Bayes' rule to

$$p_\omega(\omega|v) = \frac{p_\omega(v|\omega)p(\omega)}{p_\omega(v,\omega)}.$$

Then we have the following definition:

Definition 3.1 *The expected value of observing the platform announcement is given by the function*

$$\Delta(m|\widetilde{e}) = \int [U(\mathbf{x}^*(m, a^*(\omega))|a^*(\omega)) - U(\widetilde{\mathbf{x}}(m)|a^*(\omega))]p_\omega(\omega|v)d\omega.$$

We do not need to consider $Z(a|m)$ at this stage because the private value of information on a is independent of agents' preferences over the public policy. Notice that for each given realization of ω we will have a different ex post value of making an informed private choice. But since the actual realization of ω is ex ante unknown, the ex ante value of information must be expressed in expected terms over ω.

It is then possible to prove the following:

Proposition 3.1 *Assume $U(.) \in \Re_+$ is quasi-concave and homogeneous of degree 1 in* $\mathbf{x}$. *Then $\frac{\partial\Delta(.,e)}{\partial m} > 0$.*

Proof. See Appendix.

Note that this result can be proved whether $a^*(\omega)$ is a continuous or a discrete function. The only reason we are working with a continuous framework is to stress the fact that each agent's probability of being pivotal is zero. However, all the results are still valid with a finite number of citizens (and therefore a discrete $a^*(\omega)$) as long as we assume that the probability of being pivotal in the election is negligible (see Appendix).

Having derived an expression for the value of information, we are now ready to turn to the effort allocation problem. Let us remove the assumption that $e = \widetilde{e}$ and write the

problem of a generic agent as[16]:

$$\max_{e\in\mathcal{E}}[q(e)\int U(\mathbf{x}^*(m(1-\pi e),a^*(\omega))|a^*(\omega))p_\omega(\omega|v)d\omega$$
$$+(1-q(e))\int U(\widetilde{\mathbf{x}}(m(1-\pi e))|a^*(\omega))p_\omega(\omega|v)d\omega - ve]$$

where $q(e)$ is the probability of observing platforms. Using definition 3.1 the problem can be re-written as

$$\max_{e\in\mathcal{E}}[\int U(\widetilde{\mathbf{x}}(m(1-\pi e))|a^*(\omega))p_\omega(\omega|v)d\omega + q(e)\Delta(m,e) - ve] \tag{3.6}$$

Note that, by Assumption 3.1, we have that

$$\begin{aligned}\mathbf{x}^*(m(1-\pi e),a^*(\omega)) &= [m(1-\pi e)]\mathbf{x}^*(a^*(\omega))\\ \widetilde{\mathbf{x}}(m(1-\pi e)) &= [m(1-\pi e)]\widetilde{\mathbf{x}}\end{aligned}$$

and therefore we get

$$\begin{aligned}U(\mathbf{x}^*(m(1-\pi e),a^*(\omega))|a^*(\omega)) &= [m(1-\pi e)]V^*(\omega)\\ U(\widetilde{\mathbf{x}}(m(1-\pi e))|a^*(\omega)) &= [m(1-\pi e)]\widetilde{V}(\omega)\end{aligned}$$

To simplify notation, let us also define the following quantities:

$$\begin{aligned}V^* &= \int V^*(\omega)p_\omega(\omega|v)d\omega\\ \widetilde{V} &= \int \widetilde{V}(\omega)p_\omega(\omega|v)d\omega\\ \Delta^* &= V^* - \widetilde{V}\end{aligned}$$

This means the value of information can be written as

$$\Delta(m,e) = [m(1-\pi e)]\Delta^*$$

[16]Note that in the effort allocation problem we neglect the fact that possible monetary costs of information gathering change the endowment of voters and might therefore change their preferences over a. This is a second order effect and clearly a negligible one.

Therefore, the maximization problem (3.6) can be re-written as

$$\max_{e \in E} [m(1-\pi e)]\widetilde{V} + q(e)[m(1-\pi e)]\Delta^* - ve \tag{3.7}$$

Solving this problem, we obtain the optimal effort function $e^*(m,v)$ (remember that agents are heterogeneous in m and v). This then gives the probability of being informed on political platforms $Q(m,v|\omega) = q(e^*(m,v))$, where conditioning on ω indicates that there is one such function for each realization of ω. In particular, to link the probability of being informed to policy preferences, it is essential to understand how effort choice is dependent on the initial endowment of agents and therefore to calculate $\frac{de^*(m,v)}{dm}$.

Proposition 3.2 *If Assumption 3.1* is satisfied *then* $\frac{de^*(m,v)}{dm} > 0$ *and therefore the probability of being informed on political platforms* $Q(m,v|\omega)$ *is such that* $Q'_m > 0$.

Proof.: *see Appendix.*

Before concluding this section, let us recall that we are dealing with the private value of information; however, since the number of citizens is very large (it is actually infinite) any incentive to acquire information for political purposes (i.e. for instrumental voting) is negligible, in the sense that the probability of being a pivotal voter is zero in a continuum of agents. Therefore $Q(m,v|\omega)$ fully represents the probability each citizen has of being informed on political platforms.

3.4 Voting decisions and political competition

In this section we analyse the political competition game and citizens' private and public decisions. We will solve the game backward, deriving agents' best responses and then the political equilibrium.

3.4.1 Consequences

As we have full commitment to platforms, the policy proposed by the winning party (a^*) is implemented after the election; if the two parties get an equal share of votes then each policy is implemented with probability equal to $\frac{1}{2}$. Note that the population of voters consists of those agents who actually vote, and is therefore a subset of the entire population.

At the end of this period the realized utility for each agent will be given by

$$U(\mathbf{x}^*(m,a^*)|a^*) + Z(a^*|m) - ve^*(m;v) \tag{3.8}$$

if informed and

$$U(\widetilde{\mathbf{x}}(m)|a^*) + Z(a^*|m) - ve^*(m;v) \tag{3.9}$$

if uninformed.

3.4.2 Voting and private decisions

There are only two parties in the model, therefore strategic voting is equivalent to sincere voting. Voters always have a weakly dominant strategy and their optimal voting strategy $i^*(m,a_L,a_R)$ can have a simple representation:

$$i^*(m,a_L,a_R) = \left\{ \begin{array}{c} L \text{ if } W(a_L,m) - W(a_R,m) > 0 \\ R \text{ if } W(a_L,m) - W(a_R,m) < 0 \\ \text{abstain if } W(a_L,m) - W(a_R,m) = 0 \end{array} \right\} \tag{3.10}$$

Voters who do not observe the platforms are indifferent between the two parties and we will assume they abstain. It is important to notice the crucial difference between the behaviour of the informed, who can make their choice contingent on $\{a_L,a_R\}$, and that of the uninformed, who cannot. Thus, uninformed voters cannot be responsive to different platform announcements. The assumption that the uninformed abstain is not essential, and it only helps in providing an empirical interpretation of our results. An alternative could be to assume that they vote randomly: this would not have any impact on our results.

Optimal private decisions will be

$$\mathbf{x}^* = argmax_{x\in X} U(\mathbf{x}|a^*) \tag{3.11}$$

for informed agents and

$$\widetilde{\mathbf{x}} = argmax_{x\in X} \int U(\mathbf{x}|a^*(\omega))p(\omega|v)da \tag{3.12}$$

for the uninformed. These private decisions are made before elections[17]. Hence, when making private choices citizens do not know the election outcome, although they can form rational expectations.

3.4.3 Information gathering

At this stage we have the process described in the previous section. Agents must decide how much effort to devote to information gathering. Solving the maximization problem (3.7) we derive the optimal effort of each citizen $e^*(m, v)$ and then the probability of being informed on platform announcement $Q(m, v|\omega)$. At the end of this period the total population will be divided into informed agents (those who observe the platforms) and uninformed. Note again that more effort only implies a higher probability of being informed.

3.4.4 Platforms' announcement and political equilibrium

Parties announce their platforms simultaneously. Remember that at the beginning of the game they both observed the realization of the random variable ω and therefore they know

$$E_\omega(e|m) = \int e^*(m, v)p(v|\omega)dv. \tag{3.13}$$

It is impossible to know ex ante who is going to be informed and who is not, because this depends on the realization of the idiosyncratic shocks and because $q(e)$ represents only a probability of getting information[18]. What the parties can do is to exploit the ex ante information on observables (m and ω) and their relationship with the probability of being informed. ε is an idiosyncratic shock with no systematic relation with policy preferences, so it is irrelevant for parties' strategies[19]. Therefore, from the point of view of the parties

[17] For our purposes they could also be simultaneous to elections.

[18] It should be noted that the process of information gathering is considered ex ante. In other words, if a lucky agent observes the announcement immediately he will stop putting effort into information gathering, before reaching the ex ante optimal level e^*. However, this interim process is not observable for the parties, which can look at the situation only from an ex ante perspective. Moreover, since luck does not depend on policy preferences, this consideration will be irrelevant when coming to political proposals. This is the same argument we use for ε, which is both uncorrelated with policy preferences and unobservable for parties.

[19] Moreover we assume that only ω is observed.

we can consider

$$Q_\omega(m) = \int Q(m, v|\omega)p(v|\omega)dv. \tag{3.14}$$

We assume parties are interested in maximizing expected plurality $P^E(a_i, a_j) = E[n_i(a_i, a_j) - n_j(a_i, a_j)]$. Therefore the problem of party i $(i = L, R)$ is

$$\max_{a_i \in \mathcal{A}} P_i^E(a_i, a_j) = \int_{\mathcal{A}_i(a_i, a_j)} Q_\omega(z^{-1}(a))y(a)da - \int_{\mathcal{A}_j(a_j, a_i)} Q_\omega(z^{-1}(a))y(a)da \tag{3.15}$$

where $\mathcal{A}_i(a_i, a_j)$ represents the set of policies preferred by citizens choosing party i, given that platforms are (a_i, a_j). A Nash equilibrium in platforms (a_i^*, a_j^*) must therefore satisfy

$$P_i^E(a_i^*, a_j) \geq P_i^E(a_i^*, a_j^*) \geq P_i^E(a_i, a_j^*), \ i, j = L, R \tag{3.16}$$

Notice that from Assumption 3.2 the policy space admits a Condorcet winner[20]. When we say that a policy space admits a Condorcet winner we basically assume that everybody in the population space is capable of choosing his or her preferred option in a pairwise comparison. This is clearly not possible if some agents do not know what the available options are. However, we can still find a Condorcet winner given that any subset of the population satisfies Assumption 3.2. Given our assumptions, the Condorcet winner is the platform preferred by the voter who is median in the set of the ex post informed voters N_I. Parties clearly do not know the identity of informed and uninformed citizens and therefore cannot say ex ante what is the relevant set of voters. Since the population is very large and since both the preferred policy and the probability of being informed are monotonically related to income, we can find a focal point for parties' strategies. The relevant set of voters is ex ante an unknown set; hence the parties maximize over the expected relevant set of voters. Thus, the likelihood of being informed may be taken into account in maximizing expected votes, and this is reflected in the payoff function in the (3.15).

[20] See Gans and Smart (1996).

3.4.5 Characterization of equilibrium

In this section we derive some important properties of the equilibrium. An equilibrium in this game is given by a platform announcement for each party

$$a_i^*(\omega) \quad (i \in \{L, R\}),$$

a vector of decision strategies for informed citizens

$$\{e^*(m; v),\ i^*(m, a_L, a_R),\ \mathbf{x}^*(m, a_L, a_R)\}$$

and one for uninformed citizens

$$\{e^*(m; v),\ i^*(m),\ \mathbf{x}^*(m)\}$$

We are interested in the political equilibria, and so we leave in the background the equilibrium in private choices, which will not affect our results.

The existence of a "weighted Condorcet winner", and therefore competition among parties to reach it, ensures that political equilibrium will have some simple and intuitive properties.

Proposition 3.3 The unique political equilibrium is given by a^* s.t.

$$\int_{\underline{a}}^{a^*} Q_\omega(z^{-1}(a))y(a)da = \int_{a^*}^{\overline{a}} Q_\omega(z^{-1}(a))y(a)da.$$

Proof. See Appendix

Hence, parties will converge on the platform preferred by the expected median informed voter. The argument for this convergence is identical to the standard Downsian one, the only difference being that the relevant population distribution is weighted by the probability each citizen has of being reactive to political proposals.

Given the continuity of the policy space and of the distribution function of the cost of information, we can also prove the following result, which was used (but not proved) in the previous section.

Lemma 3.1 *The political equilibrium of this game can be expressed as a continuous*

function $a^*(\omega) : S_\omega \to \mathcal{A}$.

Proof. See Appendix.

We can now turn back to the issue of the value of information. In Proposition 3.1 we proved that the value of political information is increasing in each agent's income; in Lemma 3.2 we show that the value of information is positive, even if agents are able to understand they are in a political equilibrium: rational expectations rule out all policies that cannot be sustained in equilibrium, whatever the realization of random variables, but agents are still uncertain about which equilibrium they are in.

Lemma 3.2 *In equilibrium the value of information on platforms is positive.*

Proof. Since the distribution $Q_\omega(z^{-1}(a))$ depends on the realized value of ω, voters, who have rational expectations but do not know ω, will expect to have in equilibrium $a^*(\omega)$. Anyway, informed voters can fully deduce a^* from platform convergence. Uninformed voters rationally rule out any other possibility apart from $a^*(\omega)$ but are still uncertain about the actual a^* . This fact gives a positive value to information about parties' platforms.■

3.5 Implications for income redistribution

We can now turn to redistributive policies. As we noted earlier, little empirical support has been found for positive models of income redistribution that are based on the median voter theorem: in general, redistributive policies do not appear to be very responsive to the median/mean income ratio.

It is clear that many issues are at stake in democracies and that there is no simple way to explain redistribution. However, in this section we want to ask if information on politics may give some insights even in a simple one-dimensional framework.

It is possible to characterize the equilibrium in terms of the policy outcome in a precise way and compare it with the outcome of a standard Downsian model with perfect information.

Proposition 3.4 *Let us indicate with* a_M^* *the political equilibrium when the entire population is informed on platform announcements. Then*

$\omega \geq 0 \Rightarrow a^*(\omega) \leq a^*(0) < a_M^*$.

Moreover, $\frac{\partial^2 Q(.,.)}{\partial \omega \partial a} < 0 \Rightarrow a_\omega^{*\prime} < 0$.

Proof. See Appendix.

Political equilibrium in our game involves a public policy that will be, in general, different from that preferred by the median voter over the entire population (see Figure 3-2). The weight attached to agents by political parties is increasing in their income, and therefore the pivotal voter has an income higher than the median[21]. As long as acquiring information has a cost, the public policy will be bounded above by $a^*(0)$, which is lower than the median voter outcome. This provides a microfoundation for the idea that richer agents have more power in the political process. This is not a new idea in political science and political economy: it is for example a crucial assumption in Bebabou (2000). Nevertheless, microfounded justifications for such hypothesis are still missing in the literature.

Abusing of this result and interpreting non-responsiveness to policies as abstention in general elections, we can link this idea to the stylized facts that abstention is more common among low income agents and that countries with higher turnout tend to have higher levels of social expenditure. Starting with the classical study of Wolfinger and Rosenstone (1980), a vast empirical literature consistently finds positive correlations between turnout and variables like income or education. Some theoretical research has linked information to participation. In decision-theory terms, being better informed allows better choices and therefore should increase the probability of voting (Matsusaka, 1995). When strategic interactions are considered, less informed citizens might abstain in order to increase the probability of the better informed being pivotal (Feddersen and Pesendorfer, 1996). This, however, is only true if citizens' preferences are not too heterogeneous (Caillaud and Tirole, 1997).

In terms of our model, if we introduce a cost of voting that is independent of policy preferences, then we can easily link our results on rational ignorance to actual voter turnout[22]. This would deliver observable conclusions about electoral participation and social spending. Interestingly, Lindert (1996) finds evidence of this: "a stronger voter

[21] Analogously, Stromberg (2002) finds that office-seeking candidates will bias their policy proposals in favour of citizens that are more likely to be targeted by the mass media.

[22] Provided we have that the probability to be pivotal is non-zero.

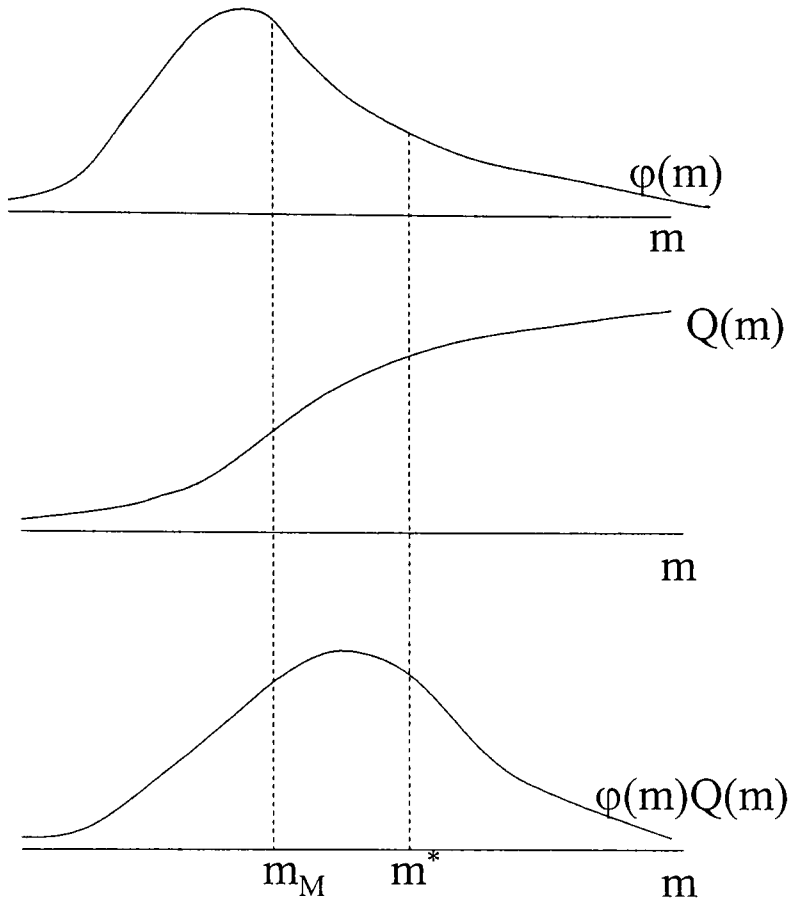


Figure 3-2: The "weighted" median voter

turnout seems to have raised spending on every kind of social program, as one would expect if one assumed that the social programs cater to the lower income groups whose voter turnout differs most over time and across countries". Section 1.5 of this thesis discusses a number of other works presenting evidence in that sense[23].

Another important conclusion of the analysis of redistributive policies within the Downsian framework is that an increase in income inequality (measured as the ratio between the mean and the median income) should lead to more redistribution. In comparing two income distributions φ_1 and φ_2 with the same mean, a way to say that φ_2

[23]However, on a study on US panel data, Besley and Case (2002) find that turnout has little effect on the party composition of legislature.

induces more redistribution than φ_1 is

$$\int_{\underline{m}}^{m_1} \varphi_2(m)dm > \frac{1}{2} \tag{3.17}$$

where m_1 is the income of the pivotal voter under distribution φ_1. The reason that the change produces more redistribution is that the pivotal voter under φ_2 is poorer (being $m_1 > m_2$, with m_2 s.t. $\int_{\underline{m}}^{m_2} \varphi_2(m)dm = \frac{1}{2}$) and therefore his or her distance from the mean has increased.

In our model, however, the condition for more redistribution translates into

$$\int_{\underline{m}}^{m_1^*} Q(m)\varphi_2(m)dm > \frac{\int_{\underline{m}}^{\overline{m}} Q(m)\varphi_2(m)dm}{2} \tag{3.18}$$

where m_1^* is the income of the pivotal voter in the distribution $Q(m)\varphi_1(m)$. It is clear that condition (3.17) does not imply condition (3.18) nor the vice versa. In general, the foregoing analysis leads to a result of indeterminacy. A mean-median ratio increase does not necessarily lead to more redistribution in a democratic system, as this will have two contrasting effects: more inequality increases the middle classes' desire for redistribution, but it also means greater dispersion in the probability of being informed, resulting in parties targeting higher-income groups. Unfortunately, it is not possible to characterize the effects of an increase in inequality better, if not in obscure and not very useful ways. However, this indeterminacy should at least counsel more prudent use of voting models for comparing the redistributive outcomes of different degrees of inequality. We can summarize this negative result in the following proposition:

Proposition 3.5 *An increase in the mean-median income ratio is neither necessary nor sufficient for more redistribution.*

In focusing on the formal aspects of electoral processes, the voting literature seems to have neglected the role of factors that certainly matter for the proper functioning of democracy: democratic decisions require not only that people go to the polls but also other institutional elements, such as those that foster informed public opinion. New elements come to play a role in our analysis. First of all the shape of the function $Q(m)$ matters. Since the results are driven by the fact that Q'_m is positive, it can be argued that traditional results are likely to be reversed when Q'_m is large enough. That is, to be able to

say anything about redistribution we must also be able to determine the impact of income on the decision to acquire information. Clearly, this may depend on many elements: a sufficiently high general level of education, for example, is an important condition for widespread access to information and certainly raises the capability to extract information from the new (we addressed this issue in chapter 2). Also, the importance of the supply of information should not be underestimated: sufficiently free press and competition in the information market can increase the availability of good quality information and hence increase political knowledge. Evidence from the previous chapter support this type of claims.

Another consideration is that focusing on median and mean incomes can be highly misleading. It would be more appropriate to consider the whole income distribution, since the identity of the expected pivotal voter can be modified by changes outside the median-mean range: changes in the distribution that leave both median and mean incomes unaltered may nevertheless influence policy choices by affecting citizens' responsiveness in other parts of the distribution, thus changing the identity of the pivotal voter.

This leads to another consideration, namely that not only relative but also absolute inequality matters. Two distributions with the same degree of relative inequality (as gauged for example by Lorenz curves) may produce different political outcomes, because the function $Q(m)$ is not necessarily linear, and will therefore "weight" the two distributive profiles differently. In section 3.3 we derived results on $Q'_m(m)$, but nothing general can be said about $Q''_m(m)$. That is, a change in the difference between mean and median income, leaving their ratio unaffected, would change the political equilibrium in our model even when it would not affect a standard Downsian model.

Furthermore the mean-median ratio (or distance) is not necessarily a good measure of inequality[24]. Indeed, we can think of an increase in inequality (in terms of Lorenz dominance, for example) associated with a reduction of the distance between mean and median income. However, as political equilibria have been derived in the literature in terms of this measure, it has become standard to consider only mean and median income. Yet our analysis suggests the need to considering the entire distribution. Further analysis is necessary to derive results in this direction.

[24] For example it does not satisfy the Pigou-Dalton transfer principle. See Lambert (1995).

Since Proposition 3.5 is essentially a negative result, we now use two examples to illustrate the possible implications of the foregoing analysis.

Example 1 (A poor majority).

Let us consider a population divided into two groups, *Poor* and *Rich*, with respective income m_P and m_R and $m_R > m_P$; we also assume that $N_P > N_R$. The two sources of information cost ω and ε now assume a finite number of possible values; in particular $S_\omega = \{\omega_L, \omega_H\}$ (with $\omega_L < \omega_H$) and $S_\varepsilon = \{\varepsilon_L, \varepsilon_M, \varepsilon_H\}$ with $(\varepsilon_L < \varepsilon_M < \varepsilon_H)$ and the respective probabilities are p_H, $p_L = 1 - p_H$, q_H, q_M, q_L. We then have the following possible realizations for the cost of information v :

$$v = \left\{ \begin{array}{c} v_H = \omega_H + \varepsilon_H \text{ w.p. } p_H \times q_H \\ v_{MH} = \omega_H + \varepsilon_M = \omega_L + \varepsilon_H \text{ w.p. } p_H \times q_M + q_H \times (1 - p_H) \\ v_{ML} = \omega_H + \varepsilon_L = \omega_L + \varepsilon_M \text{ w.p. } p_H \times q_L + q_M \times (1 - p_H) \\ v_L = \omega_L + \varepsilon_L \text{ w.p. } (1 - p_H) q_L \end{array} \right\}$$

Moreover, the probability of being informed assumes an extreme form:

$$q(e) = \left\{ \begin{array}{l} 1 \text{ if } e > \widehat{e} \\ 0 \text{ if } e < \widehat{e} \end{array} \right\}$$

We will also assume that the value of information and the income distribution are such that at a cost v_H nobody is informed, at a cost v_{MH} only the rich buy information, i.e.

$$\begin{aligned} e^*(m_R, v_{MH}) &= \widehat{e} \\ e^*(m_P, v_{MH}) &= 0 \end{aligned}$$

and at cost v_{ML} and v_L all agents value information on political party platforms at more than the cost of acquiring it, i.e.

$$e^*(m_P, v_{ML}) = \widehat{e}$$

It is immediately clear that with full information the Condorcet winner is the policy preferred by the poor $a^* = a_P$. Let us now analyse imperfect information. Using Bayes'

rule, after observing his or her own private cost, each agent is able to deduce that

$$\begin{aligned}
\Pr(\omega_H|v_H) &= 1 \\
\Pr(\omega_H|v_{MH}) &= \overline{p} = \frac{p_H \times q_M}{p_H \times q_M + q_H \times (1-p_H)} \\
\Pr(\omega_H|v_{ML}) &= \underline{p} = \frac{p_H \times q_L}{p_H \times q_L) + q_M \times (1-p_H)} \\
\Pr(\omega_H|v_L) &= 0
\end{aligned}$$

Let us then consider the two possible realizations of ω.

Case 1: $\omega = \omega_H$. Some agents will have a private cost v_H and will have no incentive to gather information. The rich with costs v_{MH} and v_{ML} will gather information. For a large population, each agent's probability of being informed can be translated into the fraction of the population that is informed. Therefore we have $(q_M + q_L)N_R$ informed. In the same way we have $q_L N_P$ informed. If $(q_M + q_L)N_R > q_L N_P$ then $a^* = a_R$.

Case 2: $\omega = \omega_L$. The poor with $v = v_{MH}$ stay uninformed while those with $v = v_{ML}$ have a value of information greater than its cost and therefore acquire it. The rich with v_{MH} and v_{ML} will acquire information. Let us now assume that $N_R < q_M N_P$. What happens to agents with cost v_L? Notice that those agents would receive a positive value from acquiring information on party platforms. However, they also have degenerate beliefs on the realization of ω. Knowing that $\omega = \omega_L$ they learn that a fraction q_M of the poor are informed, and that is enough to establish that $a^* = a_P$. Therefore they do not need to gather information to be informed on the policy, independently of their income, and can free ride on the group with higher cost $v = v_{ML}$. Only a fraction q_M of the poor and $(q_M + q_H)$ of the rich will be informed, which ensures $a^* = a_P$. Thus, we have proved the following proposition:

Proposition 3.6 Assume $q_M N_P > N_R$ and $(q_M + q_L)N_R > q_L N_P$. Then $a^* = a_R$ w.p. p_H and $a^* = a_P$ w.p. $(1 - p_H)$.

It is therefore clear that, depending on the parameters, even a small minority of rich people may be able to obtain their preferred policy. This is likely to happen when the majority of the poor are not in a position to make relevant private decisions. If, for example, a majority of the population is at a subsistence level of income, they might have low incentives to be informed on public policies, thus leaving public decisions to the

rich minority, in spite of the fact that collective decisions could significantly affect their welfare. When we compare this with the outcome under full information, it is evident that the probability of having outcome $a^* = a_P$ has been reduced from 1 to $(1 - p_H)$.

Example 2 (Constitutional restriction). Let us consider again a population divided into rich and poor, with the same assumptions on population distribution and the cost of information as in Example 1. We will now also see that other sources of uncertainty can be introduced and that something can be learned about them from parties' behaviour. Also, initial endowment does not need to be income.

Agents have identical utility functions $U(d, l, g)$, where d is consumption, l is leisure and g is a public good. Gross income and net income are respectively generated by agent i according to

$$\begin{aligned} m_i &= w_i(1 - l) \\ d_i &= m_i(1 - \tau) \end{aligned}$$

where w_i is the wage rate, $(1 - l)$ is labour supply (with total time normalized to 1) and τ is a flat tax rate. The public good is produced with constant returns at unitary cost and, assuming balanced budget, we have

$$g = \tau \sum m_i$$

Rich and poor are endowed with different wage rates $w_P < w_R$. That of the rich is assumed given and common knowledge, while that of the poor is a random variable that can assume two possible realizations: $w_P = \underline{w}_P$ w.p. $\underline{p}$ and $w_P = \overline{w}_P$ w.p. $(1-\underline{p})$, with $\overline{w}_P > \underline{w}_P$. Notice that nothing would change if instead of uncertainty on the wage rate we considered any element of preferences, like intensity of preference for the public good by either of the two groups.

We consider two possible regimes: in regime *(a)* a linear tax is levied on the entire population and the revenue is used to produce the public good. In regime *(b)* a constitutional restriction prevents taxation below a threshold level of gross income, so that if the poor have wage rate $\underline{w}_P$ they are not taxed, whatever the tax rate. Indicating this

threshold with $\widehat{m}$, we have

$$m_R(w_R,\tau) > \ m_P(\overline{w}_P,\tau) > \widehat{m} > m_P(\underline{w}_P,\tau) \ \forall\tau$$

Note that the public policy issue τ is unidimensional, since there is a binary correspondence between τ and g.

Case (a): with full information τ is known to everybody. Therefore each agent will perform an individual optimization over labour supply, taking into account his or her own wage rate and the tax. The indirect utility function after this process is given by $V(w(1-\tau),g)$. Since preferences are assumed identical for all agents, when coming to the public policy issue we will typically have $\tau_P > \tau_R$ (and $g_P > g_R$). Therefore the Condorcet winner is represented by $\tau^* = \tau_P$, and competing political parties will converge on τ^*. Let us now consider the case of imperfect information. This is very similar to that analysed in the previous example. Therefore, on the basis of Proposition 3.6, if we did not have uncertainty on the wage rate we could have concluded that $\tau^* = \tau_R$ w.p. p_H and $\tau^* = \tau_P$ w.p. $(1-p_H)$. But now we have to take into account that the optimal tax rate for each agent depends on the realization of the uncertain wage rate of the poor. We will have

$$\tau_P(\underline{w}_P) > \tau_P(\overline{w}_P) > \tau_R(\overline{w}_P) > \tau_R(\underline{w}_P)$$

Therefore:

$$\tau^* = \left\{ \begin{array}{c} \tau_R(\underline{w}_P) \text{ w.p. } p_H \times \underline{p} \\ \tau_R(\overline{w}_P) \text{ w.p. } p_H \times (1-\underline{p}) \\ \tau_P(\overline{w}_P) \text{ w.p. } (1-p_H) \times (1-\underline{p}) \\ \tau_P(\underline{w}_P) \text{ w.p. } (1-p_H) \times \underline{p} \end{array} \right\}$$

The value of information for each rich agent is represented by

$$\begin{aligned} \Delta(w_R) \ = \ & V(w_R(1-\tau^*),g^*) - [\ p_H \times \underline{p} V(w_R(1-\tau_R(\underline{w}_P)),g_R(\underline{w}_P)) \\ & + p_H \times (1-\underline{p}) V(w_R(1-\tau_R(\overline{w}_P)),g_R(\overline{w}_P)) \\ & + (1-p_H)\underline{p} V(w_R(1-\tau_P(\underline{w}_P)),g_P(\underline{w}_P)) \\ & + (1-p_H)(1-\underline{p}) V(w_R(1-\tau_P(\overline{w}_P)),g_P(\overline{w}_P))] \end{aligned}$$

The poor learn the realization of their own wage rate and therefore have one less source

of uncertainty. Then the value of information for each poor agent is given by

$$\begin{aligned}\Delta(w_P) &= V(w_P(1-\tau^*), g^*) - [\, p_H V(w_P(1-\tau_R(w_P)), g_R(w_P)) \\ &\quad +(1-p_H)V(w_P(1-\tau_P(w_P)), g_P(w_P))\end{aligned}$$

We are making the following assumption on the value of information:

$$\begin{aligned}v_H &> \Delta(w_R) > v_{MH} > v_{ML} > v_L \\ v_H &> v_{MH} > \Delta(w_P) > v_{ML} > v_L\end{aligned}$$

Note also that the value of information for the rich in this case is not limited to knowledge of policies but extends to knowledge of an exogenous element (the wage rate of the poor) that can be useful for some decisions and that is revealed by politicians' behaviour.

Case (b). Now we have a constitutional restriction that prevents the poor from being taxed if they are endowed with the low income. The preferred tax levels will change accordingly. Indicating with $\tau_P^C(\underline{w}_P)$ the tax rate preferred by the poor when their wage rate is low, under the constitutional restriction we have $\tau_P^C(\underline{w}_P) = 1 > \tau_P(\underline{w}_P)$. If the wage rate of the poor is high then their preferred tax rate is not affected by the constitutional restriction, so $\tau_P^C(\overline{w}_P) = \tau_P(\overline{w}_P)$.

The preferred tax rate of the rich also changes. If the poor's wage is high then again the constitutional restriction has no effect: $\tau_R^C(\overline{w}_P) = \tau_R(\overline{w}_P)$. But if $w_P = \underline{w}_P$ then $\tau_R^C(\underline{w}_P) < \tau_R(\underline{w}_P)$ (assuming that the substitution effect dominates the income effect).

With full information the constitutional restriction is clearly favourable to the poor since the new Condorcet winner will simply follow the preferences of the poor. Therefore the equilibrium policy becomes $\tau^* = 1$ w.p. $\underline{p}$ and remains $\tau_P(\overline{w}_P)$ w.p. $(1-\underline{p})$. When introducing imperfect information, notice that under the constitutional restriction if the wage rate realized for the poor is low, then they have no uncertainty over their own tax rate, which is going to be zero independently of public choice. The poor can then perform their preferred labour supply choice without information gathering: the value of information for them becomes zero and therefore lower than the lowest possible realization for the cost of information. If this is the case then the Condorcet winner is represented

by $\tau_R^C(\underline{w}_P)$. However, if the realization of the wage rate is high then the poor will still gather information and therefore the Condorcet winner is $\tau_P^C(\overline{w}_P) = \tau_P(\overline{w}_P)$, i.e. exactly the tax rate that would prevail without constitutional restriction. As a consequence, the political equilibrium is $\tau_R^C(\underline{w}_P)$ w.p. $\underline{p}$, $\tau_P(\overline{w}_P)$ w.p. $(1-\underline{p})(1-p_H)$, and $\tau_R(\overline{w}_P)$ w.p. $(1-\underline{p})p_H$.

The situation considering asymmetric information has been reversed. Now we can have a deviation from $\tau_P(\overline{w}_P)$ with probability $\underline{p} + (1-\underline{p})p_H$; not, however, towards an increased tax but a reduced one. Moreover, the constitutional restriction could be harmful for the poor. Without the constitutional restriction the (ex ante) expected tax rate is

$$\begin{aligned} E(\tau^*) &= \tau_R(\underline{w}_P)\, p_H \times \underline{p} + \tau_R(\overline{w}_P)\, p_H \times (1-\underline{p}) \\ &\quad + \tau_P(\overline{w}_P)(1-p_H) \times (1-\underline{p}) + \tau_P(\underline{w}_P)(1-p_H) \times \underline{p} \end{aligned}$$

while under the constitutional restriction we have

$$E^C(\tau^*) = \tau_R^C(\underline{w}_P)\underline{p} + \tau_P(\overline{w}_P)(1-\underline{p})(1-p_H) + \tau_R(\overline{w}_P)(1-\underline{p})p_H..$$

It is easy to verify that $E(\tau^*) > E^C(\tau^*)$, and therefore a restriction which has been introduced in order to increase income redistribution might eventually reduce it.

It is clear that a restriction on targeted benefits instead of one on the tax would have delivered the same conclusion. The basic result is that to participate in public life people may need some "selective incentives", and an important aspect of public policies is whether or not they generate such incentives.

3.6 Extensions on coalitions and opinion leaders

The solution concept used in the model is Nash equilibrium: nobody wants to deviate unilaterally from his or her best response given the behaviour of other agents. It is well known that Nash equilibria do not need to be efficient, in the sense that Pareto improvements are sometimes possible when agents are able to coordinate.

It should first be noted that in the model presented here this is not the case. If side payments among citizens are not possible, then any agent would just prefer his or her

ideal level of a to any other. Therefore, once an equilibrium has been reached, there is no way to improve the condition of one of the citizens without putting somebody else in a worse situation. It is useful to think of a as a public policy grounded in a second best environment. This creates the possibility that some public policy choices are less efficient than others, in the sense that they could be Kaldor-Pareto dominated if some form of compensation were possible. Anyway, since we limit our policy space to one dimension (a), then the conclusion must be that any outcome of the political process is Pareto efficient. This is a typical feature of all Downsian models. It is nevertheless interesting to note that in our model the political outcome is not preferred by the majority of the population, i.e. there are available alternatives that could potentially beat in pairwise comparison that selected.

Further progress could be made by recognizing that the political equilibrium of this game does not need to be coalition-proof. Nash equilibrium is concerned with the behaviour of single agents. We know that since the probability of being a pivotal voter is zero, nobody will put more effort into information gathering than what is optimal from a private perspective. However, if a large group of citizens with similar preferences can coordinate on acquiring more information, this would shift the political equilibrium towards their preferred one. This shift in political outcome could be worth the extra-effort spent in information gathering; the problem is that information above the private needs is a public good, and individuals will fail to coordinate without some specific coordinating device.

However, in a world in which it is individually costly to gather information on political platforms, it can also be too costly to coordinate people for acquiring information: moreover, there may be other reasons why people might not be willing to coordinate on information acquisition[25]. The form of coordination one can imagine is directed to reducing the costs for some groups: this is typically done by many organizations with an interest in policy choices. Another way this coordination can, at least partially, take place, is by transmitting "cheap" information. In other words, it might not be necessary to know and perfectly understand the public budget and its implications in order to make a "good" choice. If a pre-election stage is added to our model, in which people can

[25] For example because it can seriously limit individual liberties.

simply endorse parties and say "vote for R" or "vote for L", without any justification, this could change the political outcome, as long as the announcements come from people whose preferences are known[26]. We can think that a cheap message (one that can be sent and received at low cost), rather uninformative *per se,* can nevertheless serve uninformed citizens as a good signal of where the preferred policy lies[27]. The problem in this case is transferred to the "reliability" of the sources of such messages. Is it realistic to assume that people know the political preferences of other agents? It should be recognized that some agents are able to signal their preferences in some way and that many organizations are also able to establish a reputation in this sense. Trade unions, for example, are often able to coordinate people's voting decisions because of their reputation. However, the role of those organizations or opinion leaders is not necessarily to transmit information, which could well maintain the same cost, but to convey messages that can coordinate people's actions: we can think of this as a possible direction for further investigating the role of ideologies and leadership in the political process.

The fact that this coordination failure can be more pronounced among low-income citizens is consistent with good many stylized facts about voters' turnout in elections, participation in organizations, etc. Moreover, it may tell us something about the role of political organizations in democracies, and in particular about the historical differences in the way popular parties were organized compared with traditional liberal parties (i.e. parties that were formed before universal suffrage). Our analysis may provide a rationale for the strong organization and sense of the leadership typical of most popular parties: this is simply consistent with the necessity for more effective coordination.

3.7 Summary and conclusion

This chapter studies the role of citizens' demand for political information in elections and its link with redistributive policies; it provides a possible explanation for the low empirical support encountered by Downsian models of income redistribution. This is done by linking the demand for political information to voters' responsiveness to political platforms

[26] The seminal cheap-talk game is by Crawford and Sobel (1982). They consider a sender and a decision-maker who receives the sender's message rather than a population of decision-makers. See also Lupia and McCubbins (1998).

[27] See for example Grossman and Helpman (1999).

and considering that incentives to gather information may derive from its relevance for private choices. This incentive is generally asymmetric across the population, which may generate a heterogeneous degree of awareness about policies. We consider a Downsian environment with vote-seeking parties and the possibility of full commitment to proposed platforms and show that, for a wide class of utility functions, the ex ante value of political information is increasing in income and therefore, in electoral periods, richer agents have higher probability of being informed on proposed platforms. Since parties tend to target the citizens who are expected to be more responsive to their proposed platforms, the political equilibrium involves policy convergence not to the median preferred policy but to the policy preferred by the expected median informed voter. Therefore redistribution can be expected to be, in general, less than that predicted by the median voter result. Moreover, an increase in inequality will have two contrasting effects: it will increase the desire of agents with income below the mean for redistribution, but it will also generate greater dispersion in the probability of being informed, resulting in parties targeting higher-income voters. The net effect depends on many variables and cannot be determined in a simple way, as in traditional Downsian models. This is a possible explanation for the fact that greater inequality in democratic countries does not very often lead to more social spending or redistributive taxation.

Another consequence of our analysis is that to understand redistribution we should not confine our attention to relative inequality; if a large majority of the population have only a subsistence income cannot be expected to obtain their preferred policies. At the same time, some restrictions on the policy space that are apparently beneficial for the poorest segments of the population may end up reducing their incentives to participate in public life and therefore actually working against redistribution. Those perverse effects cannot be captured in models that assume perfect information.

This analysis calls for a better understanding of mechanisms and institutions that, though not being part of a formal definition of democracy, are nevertheless quite important for its functioning. If informed choices are generally better than uninformed ones, then having an informed public opinion is an important characteristic of a truly democratic system. This consideration seems to have been neglected in most of the public choice literature to date. What is done here is clearly only a partial step, and further investigation is necessary.

From a theoretical point of view this approach can be extended to different and more sophisticated models of political competition, where the effect of multidimensional policy spaces and non-commitment on platforms can be examined taking the role of information into account. Also, the link between lack of information and abstention deserves further investigation, in particular when political platforms are endogenous. Empirical investigation could also help understanding whether parties actually target more responsive voters.

3.8 Appendix: Proof of results

Proof of Proposition 3.1 (if $a^*(\omega)$ is a continuous function).

We divide the proof in 3 steps.

1) Let us consider the objective function $\int_{\underline{\omega}}^{\overline{\omega}}\{U(\mathbf{x}|a^*(\omega))p_\omega(\omega|v)d\omega$. Note that $p_\omega(\omega|v)$ is a continuous function and never changes its sign, and $a^*(\omega)$ and $U(\cdot)$ are both continuous functions. Then we can apply the weighted mean value theorem for integrals to say that $\exists\ \widehat{\omega}$ s.t.

$$\int_{\underline{\omega}}^{\overline{\omega}} U(\mathbf{x}|a^*(\omega))p_\omega(\omega|v)d\omega = U(\mathbf{x}|a^*(\widehat{\omega}))\int_{\underline{\omega}}^{\overline{\omega}} p_\omega(\omega|v)d\omega = U(\mathbf{x}|a^*(\widehat{\omega}))$$

We do not know the actual value of $\widehat{\omega}$, which depends on the concavity of $U(\cdot)$ and on σ_ω^2. But we know that the optimal decision function derived under uncertainty is the same as that derived under one of the possible deterministic functions. Then we can express the solution to the utility maximization problem as $\mathbf{x}^*(m,\mathbf{p},a^*(\widehat{\omega}))$.

2) Note that for a homogeneous of degree 1 utility function we have $\mathbf{x}^*(m,\mathbf{p}) = m\mathbf{x}^*(\mathbf{p})$ and therefore, $V(m,\mathbf{p},a^*) = mV(\mathbf{p},a^*)$. Let us define by $\widetilde{V}(m,\mathbf{p},a^*)$ the maximum utility attainable when platforms are not observed. Suppose we have a given realization $a^*(\omega')$. The indirect utility function (ex post, i.e. if a^* is observed) is thus $V(m,\mathbf{p},a^*(\omega'))$. From step 1, we can express the solution when a^* is not observed as $\mathbf{x}^*(m,\mathbf{p},a^*(\omega''))$ for some $\omega'' \in S_\omega$. Then the ex post value of information for the realization ω' is given by:

$$\begin{aligned}\overline{\Delta}(m|\omega') &= U(\mathbf{x}^*(m,\mathbf{p},a_L,a_R)|a^*(\omega')) + Z(a^*(\omega')|m)\\ &\quad -U(\mathbf{x}^*(m,\mathbf{p},a^*(\omega''))|a^*(\omega')) - Z(a^*(\omega')|m)\\ &= m[V(\mathbf{p},a^*(\omega')|a^*(\omega')) - \widetilde{V}(\mathbf{p},a^*(\omega'')|a^*(\omega'))]\end{aligned}$$

Note that

$$V(\mathbf{p},a^*(\omega')|a^*(\omega')) - \widetilde{V}(\mathbf{p},a^*(\omega'')|a^*(\omega')) \geq 0$$

with strict inequality if $\omega' \neq \omega''$ (by the definition of value function), which implies that $\frac{\partial\overline{\Delta}(m)}{\partial m} > 0$.

3) Finally

$$\Delta(m|e) = \int \overline{\Delta}(m|\omega)p_\omega(\omega|v)d\omega$$

and

$$\frac{\partial\Delta(m)}{\partial m} = \int \frac{\partial\overline{\Delta}(m|\omega)}{\partial m}p_\omega(\omega|v)d\omega.$$

The stated proposition follows from the fact that $\frac{\partial\overline{\Delta}(m)}{\partial m} > 0$.■

Proof of Proposition 3.1 (if $a^*(\omega)$ is a discrete function)

As claimed in section 3.3, Proposition 3.1 does not actually require the continuity of $a^*(\omega)$. Let us then assume a finite but very large number of citizens N and the functions $p_\omega(\omega)$ and $p_\varepsilon(\varepsilon)$ as discrete probability functions with mass respectively over $S_\omega = \{\omega_i|p_\omega(\omega_i) > 0\}\ i = 1,...k$, and $S_\varepsilon = \{\varepsilon_l|p_\varepsilon(\varepsilon_l) > 0\}\,, l = 1,...,h$. Then, maintaining all other assumptions holding, we can provide the following alternative proof.

Note that step 2 in the previous proof still applies with (indicating with $\widehat{\omega}_j$ the true realization of ω)

$$\overline{\Delta}(m|\widehat{\omega}_j) = U(\mathbf{x}^*(m,\mathbf{p},a_L,a_R)|a^*(\widehat{\omega}_j)) - \sum_{i=1}^{k} p(\omega_i)U(\mathbf{x}^*(m,\mathbf{p},a^*(\omega_i))|a^*(\widehat{\omega}_j))$$

where $Z(a|m)$ has been neglected since it obviously cancels out.

By homogeneity of $U(.)$, we derive (as in step 2) that

$$\overline{\Delta}(m|\widehat{\omega}_j) = m[V(\mathbf{p},a^*(\widehat{\omega}_j)|a^*(\widehat{\omega}_j)) - \sum_{i=1}^{k} p(\omega_i)V(\mathbf{p},a^*(\omega_i)|a^*(\widehat{\omega}_j))]$$

This can be rewritten as

$$m\sum_{i=1}^{k} p(\omega_i)[V(\mathbf{p},a^*(\widehat{\omega}_j)|a^*(\widehat{\omega}_j)) - V(\mathbf{p},a^*(\omega_i)|a^*(\widehat{\omega}_j))]$$

By the definition of maximum value function we have

$$V(\mathbf{p},a^*(\widehat{\omega}_j)|a^*(\widehat{\omega}_j)) - V(\mathbf{p},a^*(\omega_i)|a^*(\widehat{\omega}_j)) \geq 0 \forall i.$$

Then we can write

$$\overline{\Delta}(m|\widehat{\omega}_j) = \Upsilon(\widehat{\omega}_j)m$$

where

$$\Upsilon(\widehat{\omega}_j) = \sum_{i=1}^{k} p(\omega_i)[V(\mathbf{p}, a^*(\widehat{\omega}_j)|a^*(\widehat{\omega}_j)) - V(\mathbf{p}, a^*(\omega_i)|a^*(\widehat{\omega}_j))] \geq 0.$$

Finally we have

$$\Delta(m|e) = \sum_{i=1}^{k} p(\omega_i)\Upsilon(\omega_i)m$$

from which the result is proved immediately■.

Proof of Proposition 3.2 The maximization problem is

$$\max_{e \in \mathcal{E}} [m(1 - \pi e)]\widetilde{V} + q(e)[m(1 - \pi e)]\Delta^* - ve$$

The first order condition is

$$-\pi m\widetilde{V} + [q_e^{'}(e)m(1 - \pi e) - q(e)\pi m)]\Delta^* - v = 0$$

Note that the second order condition is always satisfied:

$$[q_e^{''}(e)m(1 - \pi e) - 2q_e^{'}(e)m\pi]\Delta^* < 0 \quad \forall e$$

We can then apply the implicit function theorem to the FOC to say that

$$\frac{\partial e^*(m,v)}{\partial m} = -\frac{-\pi\widetilde{V} + [q_e^{'}(e^*)(1 - \pi e^*) - q(e^*)\pi)]\Delta^*}{[q_e^{''}(e^*)m(1 - \pi e^*) - 2q_e^{'}(e^*)m\pi]\Delta^*}$$

As we have seen, the denominator is always negative, so $\frac{\partial e^*(m,v)}{\partial m} > 0$ if and only if

$$-\pi\widetilde{V} + [q_e^{'}(e^*)(1 - \pi e^*) - q(e^*)\pi)]\Delta^* > 0$$

which implies

$$e^* < \frac{q_e^{'}(e^*)\Delta^* - \pi\widetilde{V} - q(e^*)\pi\Delta^*}{q_e^{'}(e^*)\Delta^*\pi} \tag{3A.1}$$

However, notice that to satisfy the FOC it must be that

$$e^* = \frac{q_e^{'}(e^*)\Delta^* - \pi\widetilde{V} - q(e^*)\pi\Delta^*}{q_e^{'}(e^*)\Delta^*\pi} - \frac{v}{\pi\Delta^* q_e^{'}(e^*)m}$$

which means that 3A.1 is always satisfied. Therefore $\frac{\partial e^*(m,v)}{\partial m} > 0$ and $\frac{\partial Q(m,v|\omega)}{\partial m} > 0$.

■

Proof of Proposition 3.3 By assumption 3.2 we know that for any platform pair (a_i, a_j) there exists one type of agent $\widehat{a}$ who is indifferent between the two and either

$$a_k < \widehat{a} \Rightarrow W(m_k, a_i) > W(m_k, a_j) \forall a_k < \widehat{a}$$

or

$$a_k < \widehat{a} \Rightarrow W(m_k, a_i) < W(m_k, a_j) \forall a_k < \widehat{a}$$

Define $L(a') = \int_{\underline{a}}^{a'} Q_\omega(z^{-1}(a))y(a)da$ and $R(a') = \int_{a'}^{\overline{a}} Q_\omega(z^{-1}(a))y(a)da$. Now consider $a' < a^*$. If party i chooses a' then party j will maximize $P_j^E(.,.)$ by setting $a'' = a' + \iota$, for an infinitesimal ι and getting expected votes $R(a'')$. But then a' is not a best response to a'' since, by continuity of the policy space, there exist $a'' + \iota$ that increases $P_i^E(.,.)$. But this is true for any $a' < a^*$. The same argument applies for any $a' > a^*$. Therefore the unique Nash equilibrium is given by (a_i^*, a_j^*) which delivers payoffs $P_j^E(a_i^*, a_j^*) = P_i^E(a_i^*, a_j^*) = 0$.■

Proof of Lemma 3.1 Note that the distribution $Q_\omega(z^{-1}(a))$ depends on the realized value of ω; therefore parties will make platform announcements contingent on ω. From platform convergence on the expected Condorcet winner we have that the equilibrium can be expressed as $a^*(\omega)$. We want to show that $a^*(\omega)$ is also a continuous function. Let us consider the implicit function

$$\zeta(\omega, a^*) = \int_{\underline{a}}^{a^*} Q(z^{-1}(a)|\omega)y(a)da - \int_{a^*}^{\overline{a}} Q(z^{-1}(a)|\omega)y(a)da = 0. \qquad (3A.2)$$

where a^* indicates the Condorcet winner in the distribution $y(a)Q(z^{-1}(a)|\omega)$. $\zeta(\omega, a^*)$ is clearly a continuous function (as $Q_\omega(m; \dot{v})$, $p(v; \omega)$ and $y(a)$ are continuous), strictly increasing in a^* and

$$\begin{aligned} \lim_{a^* \to \underline{a}} \zeta(\omega, a^*) &< 0 \\ \lim_{a^* \to \overline{a}} \zeta(\omega, a^*) &> 0 \end{aligned}$$

Thus, applying the implicit function theorem we can say that there exists a unique and continuous function $a^*(\omega)$ defined in S_ω and having values in $\mathcal{A}$ and such that

$\zeta(\omega, a^*(\omega)) = 0 \ \forall \omega \in S_\omega$.■

Proof of Proposition 3.4 In equilibrium with full information we have

$$\int_{\underline{a}}^{a_M^*} y(a)da = \int_{a_M^*}^{\overline{a}} y(a)da = \frac{1}{2}$$

while instead when $\omega > 0$ we have

$$\int_{\underline{a}}^{a_M^*} Q(z^{-1}(a))y(a)da \geq \int_{a_M^*}^{\overline{a}} Q(z^{-1}(a))y(a)da$$

since $Q(z^{-1}(.))$ is a monotonic decreasing function. This implies a_M^* cannot be an equilibrium since $\exists \iota$ s.t. $n_i(a_M^* - \iota, a_M^*) \geq n_i(a_M^*, a_M^*)$. Note that instead $n_i(a_M^* + \iota, a_M^*) \leq n_i(a_M^*, a_M^*)$, and therefore, by single crossing in policy preferences, deviations above a_M^* are never profitable. By the same property, any subset of N will have a Condorcet winner represented by the policy a^* preferred by the median voter in the considered subset.

Now remember that $Q_\omega(m; v) = q(e^*(m; v))$. Therefore if $\omega = 0$ then $v = \varepsilon$. Thus we have $E(e|m) = \int e^*(m; \varepsilon)p(\varepsilon)d\varepsilon$. Also, $e^*(m; \varepsilon)$ and $p(\varepsilon)$ are continuous functions, which implies $E(e|m)$ is continuous. Since $m' > m \Rightarrow e^*(m'; \varepsilon) > e^*(m; \varepsilon) \ \forall \varepsilon \in S_\varepsilon$ then $\frac{\partial E(e|m)}{\partial m} > 0$ and therefore $Q_0(m)$ is increasing in m which implies that $a^*(\omega)$ has an upper bound in $a^*(0)$ which is strictly lower than a_M^*.

To prove the second part of the statement, let us reconsider $\zeta(\omega, a^*)$ from the 3A.2. From the implicit function theorem we know that

$$a_\omega^{*'} = -\frac{\frac{\partial \zeta(\omega, a^*(\omega))}{\partial \omega}}{\frac{\partial \zeta(\omega, a^*(\omega))}{\partial a^*}}$$

The denominator is clearly positive, while the sign of the nominator is ambiguous. Therefore the sign of $a_\omega^{*'}$ is opposite to that of

$$\int_{\underline{a}}^{a^*(\omega)} \frac{\partial Q(z^{-1}(a)|\omega)}{\partial \omega} y(a)da - \int_{a^*(\omega)}^{\overline{a}} \frac{\partial Q(z^{-1}(a)|\omega)}{\partial \omega} y(a)da$$

First note that $\frac{\partial Q(z^{-1}(a)|\omega)}{\partial \omega} < 0$ which implies that both integrals are negative. If $\frac{\partial^2 Q(.,.)}{\partial \omega \partial a} < 0$ then any value of $\frac{\partial Q(z^{-1}(a)|\omega)}{\partial \omega}$ in the first integral is higher than any

value of $\frac{\partial Q(z^{-1}(a)|\omega)}{\partial \omega}$ in the second one. Since the derivative is calculated in $a^*(\omega)$ then each side has a total mass of half in terms of $y(a)$. Therefore we must have

$$\int_{\underline{a}}^{a^*(\omega)} \frac{\partial Q(z^{-1}(a)|\omega)}{\partial \omega} y(a)da > \int_{a^*(\omega)}^{\overline{a}} \frac{\partial Q(z^{-1}(a)|\omega)}{\partial \omega} y(a)da$$

which implies $a_\omega^{*'} < 0$.■

Chapter 4

The Instrumental Voter Goes to the News-Agent: Information Acquisition, Election Closeness, and the Media

4.1 Introduction and related literature

The literature explaining voting and elections using the tools of rational choice theory is vast. Rational choice models assume that voting is a mean by which to achieve public policy ends: this behaviour is called instrumental voting. Citizens care about public policies and voting is the instrument to reach them, or at least to increase the probability to get what is preferred.

This theory poses some problems, including the fact that the probability to be pivotal in large elections is normally so low that it could be considered negligible in optimization processes. This criticism can be overcome if we are ready to compromise on what we intend by a rational act. In a weak sense, agents behave rationally according to their perception of the reality, that could be different from the "objective" state of facts. In the case of voting, the probability to be pivotal in a large election is clearly very low, but it is not zero. The subjective perception of the probability of casting a decisive vote does not necessarily coincide with the infinitesimal numbers that appropriate but

cumbersome calculations would deliver (see for example Uhlaner and Grofman, 1986). Moreover, voting has been described as a "low cost-low benefit" activity (see Aldrich, 1993): it is therefore possible that even small changes in this probability might have an effect on incentives to participate in elections.

If we accept this reasoning then turnout should be larger in closer elections, when the probability to cast the decisive vote is higher. Unfortunately empirical analysis does not deliver any firm conclusion. Foster (1984), after reviewing a number of studies on the closeness-turnout linkage in the US, concludes that "the perceived probability of a tied election at the state level is not a powerful or reliable factor in explaining across-state voter participation rates in presidential elections". Grofman, Collet and Griffin (1998) study on US Senate and House of Representatives elections find evidence of higher turnout among registered voters in closer contests. Other recent studies based either on aggregate data (Kunce, 2001) or on individual-level data (Matsusaka and Palda, 1999) show instead a poor relationship between closeness and turnout. Using poll data, Kunce (2001) also shows how "the extent to which pre-election perceptions matter depends directly on how one measures the likelihood of a close contest". It seems fair to say that evidence is, at best, mixed.

In this chapter we will consider another implication of instrumental theories of voting: when elections are closer then information on candidates and platforms should be more valuable as the probability for a vote to matter is higher. Although Downs (1957) himself hints at both the "paradox of voting" (low incentives to vote) and "rational ignorance" (low incentives to gather political information) as closely related consequences of instrumental voting, the second of the two paradoxes has received less attention, in particular for what concerns the predictive implications of comparative static analysis. Thus, information acquisition should be related to the probability for a voter to be pivotal. If the suppliers of political information (mass media) are aware of this, then we should expect their behaviour to be influenced by marginality. In this sense, mass media behaviour under different circumstances will give us the possibility to provide a different kind of test of theories of instrumental voting and of the role of marginality as an incentive for participation. This clearly allows us to exploit information not used so far for this purpose.

Indeed, most people seem to believe that mass media have a vast impact on citizens'

electoral choices. Politicians appear to struggle for media attention and tend to complain when they do not receive enough space on newspapers or television. Some politicians blame the media for bad electoral performances. In some countries access to television and electoral advertising during electoral campaigns are publicly regulated and even publicly funded. All this must be based on the presumption that media are effective in influencing voters' behaviour.

Studies in this sense have not delivered any conclusive evidence, both because of an objective lack of data and because of the difficulty to identify the media effects in reality. In particular, the effects of the media could potentially be of several different types. At the very minimum, the media can be seen as informing the citizens about the different available options on the political market. Some theories, however, give to mass media more power than this and see them as capable of persuading the people by shaping their policy preferences[1]. On the other side of the spectrum, it is conceivable that even the general values and principles of a community can be affected by media choices about what is worth to report, which aspects to highlight and the way news are delivered[2].

Rational choice theory, at least as long as the assumption of given preferences is maintained, seems clearly better compatible with the first hypothesis, that can be dated back to the so-called theory of "uses and gratifications" (see Blumler and McQuail (1968)). After a period of intense studies about "the media effects", this theory called for a change of perspective by asking rather what the people do with the media. According to this theory, citizens can be seen as active users of media, with preferences, expectations and demands. Following this approach a number of studies have tried to understand why and how agents acquire and process political information, and what is the role of personal characteristics (education, prior knowledge of political matters, interest etc.) in this process (Lau and Sears (1986), Ferejohn and Kuklinski (1990), Zaller (1992)).

From an economist's standpoint this calls for a formal analysis of the political information market, with a demand for information that comes from individual preferences and

[1] One way to "shape" policy preferences is, for example, by *agenda setting* (McCombs and Shaw, 1972): the media can manipulate the salience of electoral issues, thus modifying the ranking of political priorities.

[2] Early studies in the first half of last century were based on the presumption that media were extremely powerful in conditioning people's attitudes and behaviour. After a period of "minimal effects" dominance (see Chapter 1), more recent studies have stressed the *priming* effect of the media (Iyengar and Kinder, 1987), i.e. the fact that the media, apart from informing or setting the agenda, can actually deliver a criterion of choice: this can potentially be completely independent of issues and platforms.

a supply of information provided, among others, by media firms. Some research has been conducted in this direction. Matsusaka (1995) provides a Bayesian decision-theoretical model of political information demand. Chapter 2 of this thesis models information acquisition as an individual production function, considering the role of ideological beliefs and providing evidence on the linkage between information and turnout during the 1997 general election in Britain. Specific characteristics of the media industry (like economies of scale and concentration) can also be expected to matter for the way people are informed about politics. Works in this direction are Spence and Owen (1977) and Noam (1987).

As we discussed several times in this thesis, the degree of citizens' awareness of political matters may in turn be expected to affect public policy making, both on efficiency and redistributive grounds. Examples of works in this sense are Grossman and Helpman (1999) and Lohmann (1998). Chapter 3 in this thesis models the impact of incentives to acquire information on redistribution. Media's impact in elections and implications for public policies have been studied by Stromberg (2001 and 2002), Besley and Burgess (2002), and Besley and Prat (2001)[3].

The theoretical model presented in this paper builds on Stromberg's (2002) model of mass media competition. Stromberg argues that "the increasing-return-to-scale technology and advertising financing of media firms induce them to provide more news to large groups, such as tax payers and dispersed consumer interests, and groups that are valuable to advertisers". Eventually, this information bias will be taken into account by politicians when proposing electoral platforms and will therefore translate into a policy bias.

In what follows we will use again a version of the information acquisition model presented in chapter 2. Our units of analysis (called "groups" in Stromberg's model) are represented by electoral constituencies: in this way we will be able to implement a test of our predictions, as well as of some of Stromberg's results.

One of such results concerns the effect of group size on news supply. Larger groups should receive more media attention as they provide more readership and therefore more revenue. The same can be said of groups that are more valuable to advertisers (for example wealthier groups). However, in the context of our model a countervailing effect can be identified: in larger groups we should expect a more severe collective action problem.

[3] Further details on these works are given in Chapter 1.

Thus, in larger constituencies the probability to cast a decisive vote is smaller and such will be the demand for information. This "collective action effect" can potentially offset the "group size effect"; only empirical investigation can shed further light and allow us to accept or reject any theoretical result in this sense.

Information supply also depends on the newspapers' production function. Fixed costs are normally very high but it will be clear that delivery costs could also play an important role in information supply: more densely populated areas will receive more news coverage (other things equal) simply because the cost of the marginal reader is lower in such areas.

This chapter can be summarized as follows. In the next section we will present the theoretical model of information demand and supply. Political information can be demanded for a number of reasons, including instrumental voting. Thus, it will be higher in marginal constituencies. This higher demand will induce a larger supply by profit-maximizing media firms. Media's revenue per reader is represented by the price paid for the newspaper plus the amount paid by advertisers per reader. This amount is not the same for all customers and can be expected to be higher for customers that are more valuable to advertisers. The cost of producing newspapers is fixed but there is a variable delivery cost. Thus, in equilibrium, information supply is higher in marginal constituencies as well as in constituencies with richer and more concentrated electorate. About the size of the electorate we identify two effects working in opposite directions, the "group size effect" and the "collective action effect". In section 4.3 these predictions are tested for the 1997 British general election. The test consists of two parts. The first uses aggregate data and focuses on mass media behaviour. We will use data collected from a major national newspaper during the electoral campaign, as well as electoral data and the 1991 Census. The second part will focus instead on individual behaviour and use survey data from the 1997 British General Election Study. The results suggest a high degree of compatibility between our theory and the data. Section 4.4 concludes.

4.2 The model

We start by considering a polity divided into two electoral constituencies μ and o. Each constituency elects a member of parliament (MP). There are two competing parties L and R each presenting one candidate in all constituencies. MPs are elected in a first past the

post system. With obvious notation we will indicate the candidates in each constituency with L_μ, R_μ, L_o, R_o.

Suppose the two candidates in each constituency are chosen independently by parties through a process that is unknown to citizens. This process can be represented for both parties by respective distribution functions $F_R(a)$ and $F_L(a)$ (with densities $f_L(a)$ and $f_R(a)$) over the support $\mathcal{A} \subseteq \Re_+$ of candidates' types.

For simplicity we will assume that the policy space is the same as the candidates' space and, abusing notation, that utility from policy a is a. Policies are formed at the central level by the parliament of the two MPs and affect both constituencies. If a_μ is the candidate elected in constituency μ and a_o is elected in constituency o, then the central policy will be $a^* = \frac{1}{2}a_\mu + \frac{1}{2}a_o$.

The net benefit to citizen in constituency μ from electing the preferred of the two candidates $a_{\mu 1}$ and $a_{\mu 2}$ is given by

$$\begin{aligned} B(a_{\mu 1}, a_{\mu 2}|a_o) &= |(\frac{1}{2}a_o + \frac{1}{2}a_{\mu 1}) - (\frac{1}{2}a_o + \frac{1}{2}a_{\mu 2})| \\ &= \frac{1}{2}|a_{\mu 1} - a_{\mu 2}|. \end{aligned}$$

Analogously

$$B(a_{o1}, a_{o2}|a_\mu) = \frac{1}{2}|a_{o1} - a_{o2}|.$$

We also assume that μ is marginal and this is common knowledge; i.e., if we indicate with P_i $(i = \mu, o)$ the (common) prior probability that a vote will result decisive, each agent believes that $P_\mu > P_o$. We can think of this probabilities as coming from different prior beliefs about the candidates in the two constituencies. For example in constituency μ the distribution functions $F_R(a)$ and $F_L(a)$ are "more similar" than in o. However, also the population size in each constituency will clearly play a role as a larger electorate, with given priors, will reduce the probability of each single vote to be pivotal. With only two constituencies and given electorate this effect cannot be fully appreciated in the model. However, this "collective action effect" can be expected to play a role and will be considered in the empirical investigation. A trivial way to consider this effect is to write $P_i = P_i(N_i)$ where N_i is the size (in terms of electorate) of constituency i.

4.2.1 Information demand

To avoid cumbersome notation we will focus on a generic constituency. Citizens utility from voting when types are known is then $W(a_L, a_R) = P_i B(a_L, a_R)$. However, the expected utility from an informed voting choice before candidates are selected is given by

$$W^* = P_i \int \int B(a_L, a_R) dF_L(a) dF_R(a), \; i = \mu, o$$

For simplicity, and without loss of generality, here we will assume that there is no cost of voting.

We assume voters are ex ante uninformed about candidates. We will indicate the expected utility from uninformed voting as $\widetilde{W}$. We can then define the ex post utility of an informed vote versus an uninformed one as

$$\Delta(a_L, a_R) = W(a_L, a_R) - \widetilde{W}$$

Before gathering information, however, the candidate types are unknown. Thus, the ex ante utility of gathering information is

$$\Delta = \int \int [W(a_L, a_R) - \widetilde{W}] dF_L(a) dF_R(a)$$

Lemma 4.1 $\Delta = W^* - \widetilde{W} \geq 0$.

Proof.

Let us consider a generic constituency and introduce the following notation:

$$\begin{aligned} \mathcal{A}_L^2 &= \left\{ a_L, a_R \text{ s.t. } P \int \int (a_L - a_R) dF_L(a) dF_R(a) > 0 \right\} \\ \mathcal{A}_R^2 &= \left\{ a_L, a_R \text{ s.t. } P \int \int (a_L - a_R) dF_L(a) dF_R(a) < 0 \right\} \end{aligned}$$

Suppose now that $F_L(a)$ and $F_R(a)$ are s.t. candidate L is preferred, i.e.

$$\int \int (a_L - a_R) dF_L(a) dF_R(a) > 0$$

An uninformed voter in this case votes for candidate L. Her ex ante utility is

$$\begin{aligned}\widetilde{W} &= \frac{1}{2}P_i \int\int_{\mathcal{A}_L^2} (a_L - a_R) dF_L(a) dF_R(a) - \\ &\quad \frac{1}{2}P_i \int\int_{\mathcal{A}_R^2} (a_R - a_L) dF_L(a) dF_R(a)\end{aligned}$$

The ex ante (i.e. before knowing the realization of candidates) utility of an informed vote is instead

$$\begin{aligned}W^* &= \frac{1}{2}P_i \int\int_{\mathcal{A}_L^2} (a_L - a_R) dF_L(a) dF_R(a) + \\ &\quad P_i \int\int_{\mathcal{A}_R^2} (a_R - a_L) dF_L(a) dF_R(a).\end{aligned}$$

The second term in the right-hand side is positive by definition, therefore $W^* - \widetilde{W} \geq 0$.■

Political information can be demanded for a number of different purposes. Instrumental voting is just one possibility. A sense of civic duty might also play a role as this can be seen as part of being a "good citizen". As we argued in the previous chapter, political information can also be demanded to understand or forecast public policies and this in turn can be useful for better private decision-making. Finally, information can be enjoyed as a consumption good and therefore be directly included in the utility function. We represent all this "exogenous" utility from information with Λ and say that total utility from information is

$$\Phi = \Lambda + \lambda\Delta \tag{4.1}$$

Here λ is a parameter we introduce for our convenience. Instrumental voting therefore implies that $\lambda > 0$. Otherwise we should expect $\lambda = 0$, i.e. no demand for political information arising from voting decision-making. Thanks to the following result, we will be able to test $\lambda > 0$ versus an alternative of $\lambda = 0$.

Proposition 4.1 If $\lambda > 0$ then Φ is higher in constituency μ.

Proof. Straightforward from the (4.1), as $\Delta = \Delta(P)$ with $\frac{\partial \Delta(P)}{\partial P} > 0$ and P is inversely related to expected margins of victory.

4.2.2 Information supply and mass media

We consider two newspapers Γ and Θ. They supply political news about both constituencies. We assume they have a fixed space $\overline{s}$ to devote to these news and indicate with $s_{\mu}^{\Gamma} \in [0, \overline{s}\ ]$ the space devoted by newspaper Γ to news about μ; analogously we can define $s_{o}^{\Gamma}, s_{\mu}^{\Theta}, s_{o}^{\Theta}$. Each citizen buy one newspaper. The probability for a citizen that buys newspaper j to get informed about constituency i is $q(s_i^j)$, with $q^{'} \geq 0$ and $q^{''} \leq 0$. We will assume each citizen will only care about her own constituency, thus simply ignoring news about the other constituency.

We then have $s_{\mu}^{j} + s_{o}^{j} = \overline{s}$ $(j = \Gamma, \Theta)$ and define a newspaper news profile as $\left\{ s_{\mu}^{j}, s_{o}^{j} \right\}$. A citizen living in constituency μ gets from newspaper Γ a utility from news equal to

$$\Psi(s_{\mu}^{\Gamma}) = q(s_{\mu}^{\Gamma})\Phi_{\mu}.$$

Newspapers also report about other things apart from politics. Culture, sport, and other events are also covered as well as enjoyed by readers. Each paper has its own mix over these different forms of entertainment and also its own way of dealing with them. Also, the way politics in itself can be reported is not unique. The depth and the focus of news, as well as possible partizanship, all matter for the reader. We will therefore indicate the expected utility from newspaper Γ (Θ) to citizen k in constituency μ with $\Psi(s_{\mu}^{\Gamma}) + \gamma_k$ $(\Psi(s_{\mu}^{\Theta}) + \theta_k)$, where γ_k (θ_k) is a fixed characteristic of newspaper Γ (Θ) that makes it different from Θ (Γ). Analogously for the other constituency. We are then assuming that editorial choices, entertainment content, partizanship etc. are fixed characteristics of each newspaper: this is not an unrealistic assumption in the short run and certainly within the space of an electoral campaign.

Then we say that citizen k in constituency μ buys newspaper Γ if

$$\Psi(s_{\mu}^{\Gamma}) + \gamma_k \geq \Psi(s_{\mu}^{\Theta}) + \theta_k \tag{4.2}$$

and buys newspaper Θ otherwise. Let us indicate with $\widehat{\Psi}_{\mu}$ the difference $\Psi(s_{\mu}^{\Gamma}) - \Psi(s_{\mu}^{\Theta})$ and with η_k the difference $\theta_k - \gamma_k$.

Newspapers are uncertain about individual preferences, in particular preferences about the entertainment component. We assume η_k is distributed according to a distribution

function H_i $(i = \mu, o)$, which is common knowledge. The corresponding density function is h_i. Thus, the probability that citizen k in constituency μ buys newspaper Γ is given by $Pr[\eta_k \leq \widehat{\Psi}_\mu] = H_i(\widehat{\Psi}_\mu)$.

We then introduce the following assumption, that will ensure that the pay-off functions of the newspapers are concave. This is an adaptation of condition C1 in Lindbeck and Weibull (1987).

Assumption 4.1 $\frac{|h_i'(\widehat{\Psi}_i)|}{h_i(\widehat{\Psi}_i)} \leq \frac{|q_i''(s_i^j)|}{\Phi_i(q_i'(s_i^j))^2}$, $i = \mu, o$; $j = \Gamma, \Theta$

Newspapers maximize expected profits. Each reader provides the newspaper with a revenue ρ which is the sum of the price directly paid by readers to buy the paper and the amount paid by advertisers per reader. Therefore total profits in the industry are given by $\Pi = n\rho - 2\overline{C}$, where n is the total number of citizens in the polity and $\overline{C}$ the fixed cost to produce each newspaper. For the moment, we only consider fixed costs and assume marginal costs are zero. Of course there are marginal costs represented by the printing and delivery costs but the "cost of the first copy" is normally by far the biggest cost.

Since we are interested in the share of the market newspapers have in each constituency we will rewrite the expected profit equation for newspaper Γ as

$$E(\Pi^\Gamma) = \rho[E(n_\mu^\Gamma) + E(n_o^\Gamma)] - \overline{C}$$

where n_i^j is the number of readers newspaper j has in constituency i. For newspaper Θ we have $\Pi^\Theta = \Pi - \Pi^\Gamma$. Since costs are sunk, newspapers are only interested in maximizing revenue: in our model this implies that newspapers will maximize the expected number of readers n. Indicating with N_i the total number of voters in constituency i, we will have

$$E(n^\Gamma) = N_\mu H_\mu(\widehat{\Psi}_\mu) + N_o H_o(\widehat{\Psi}_o)$$

A strategy for newspaper j is given by $s^j = [s_\mu^j, s_o^j]$. We will indicate the set of feasible strategies for newspaper j with $\Sigma_j = \{s_\mu^j, s_o^j |\ s_\mu^j + s_o^j = \overline{s}\}$.

This is a zero-sum game. Therefore a Nash equilibrium of the maximizing readership game is given by a feasible strategy profile $\{s_\Gamma^*, s_\Theta^*\}$ s.t.

$$E(n^\Gamma | s_\Gamma^*, s_\Theta) \geq E(n^\Gamma | s_\Gamma^*, s_\Theta^*) \geq E(n^\Gamma | s_\Gamma, s_\Theta^*)$$

Proposition 4.2 *Suppose that Assumption 4.1 is satisfied, $\lambda > 0$, and $N_\mu = N_o$. Then an equilibrium strategy profile $\{s^*_\Gamma, s^*_\Theta\}$ must satisfy $s^\Gamma_\mu = s^\Theta_\mu > s^\Gamma_o = s^\Theta_o$.*

Proof. The best reply function for newspaper j is defined implicitly by the first order conditions

$$\begin{aligned} N_\mu h_\mu(\widehat{\Psi}_\mu)\Phi_\mu q'(s^j_\mu) &= \vartheta \\ N_o h_o(\widehat{\Psi}_o)\Phi_o q'(s^j_o) &= \vartheta, \\ j &= \Gamma, \Theta \end{aligned}$$

where ϑ is the Lagrange multiplier associated with the problem. This implies that

$$\begin{aligned} q'(s^\Gamma_\mu) &= q'(s^\Theta_\mu) \\ q'(s^\Gamma_o) &= q'(s^\Theta_o) \end{aligned}$$

and therefore

$$\begin{aligned} s^\Gamma_\mu &= s^\Theta_\mu \\ s^\Gamma_o &= s^\Theta_o \end{aligned}$$

Now remember that

$$\begin{aligned} \Phi_\mu &= \Lambda + \lambda\Delta(P_\mu) \\ \Phi_o &= \Lambda + \lambda\Delta(P_o) \\ \frac{\partial\Delta(P_i)}{\partial P_i} &> 0,\ i = \mu, o. \end{aligned}$$

Being $\Phi_\mu > \Phi_o$ from the first order conditions we get that $s^j_\mu > s^j_o$, $j = \Gamma, \Theta$.

Assumption 4.1 guarantees that the second order conditions are satisfied (see Lindbeck and Weibull, 1987).■

So far we only focused on the implications of marginality on information demand and supply. There are a number of other factors that can have an influence on information demand and supply and therefore should be used as control variables when trying to assess the effects of election closeness. On the media revenue side it is quite realistic to assume

that not everyone has the same value to advertisers and that newspapers are capable of discriminating among different readers. The extent of this discrimination depends on the knowledge newspapers and advertisers have of market conditions and people's characteristics. Thus, we should expect this type of discrimination to become more and more relevant as new technologies improve the amount and quality of information on customers. Stromberg (2002) relates the value to advertisers to an interest in specific aspects of public policy: for example, being interested in a particular public service rather than others reveals something about people's income, and readers' income is important for advertisers. At the same time in practice we do not observe any price discrimination across different readers. This means that discrimination will mainly occur through information supply.

Another consideration concerns costs. So far we assumed that the marginal cost of producing and delivering papers was zero. Although, as we said, marginal costs have only a minor part in the production of newspapers, for our purposes delivery costs could be important. We are considering possible spatial discrimination by newspapers and in this sense delivery costs could show substantial variation. In particular, in areas which are densely populated, marginal delivery costs are probably negligible while they could be sizeable if our newspapers wanted to reach readers in remote parts of the country.

By modifying our assumptions and introducing differentiated constituencies we will therefore obtain a rationale for control variables that will make our test more reliable. At the same time in this way we will also be able to implement a direct test of some of the main Stromberg's results.

Heterogeneity here enters at the constituency level. In other terms newspapers are not able to discriminate readers according to any other individual characteristics apart from the constituency they come from, and we now assume constituencies are statistically different. This is actually the strategy that will be used to implement the empirical analysis .

Assumption 4.2 $\rho_\mu \neq \rho_o$.

Advertisers will induce from the constituency a number of other characteristics of interest and therefore will be willing to pay differently for marginal readers coming from different constituencies. Also the cost function is now different.

Assumption 4.3 *The newspaper cost function is* $TC = \overline{C} + N_\mu H_\mu(\widehat{\Psi}_\mu)\upsilon_\mu + N_o H_o(\widehat{\Psi}_o)\upsilon_o$, *where* υ_μ *and* υ_o *are the cost of the marginal reader.*

For empirical purposes we will mainly identify υ_μ and υ_o with delivery costs.

Now we can define the net marginal revenue per-reader as

$$\widetilde{\rho}_i = \rho_i - \upsilon_i, \ i = \mu, o$$

The profit equation for newspaper j can be re-written as

$$E(\Pi^j) = \widetilde{\rho}_\mu E(n_\mu^\Gamma) + \widetilde{\rho}_o E(n_o^\Gamma) - \overline{C}, \ j = \Gamma, \Theta$$

To ensure that every citizen buys one newspaper and newspapers have an interest in reaching all citizens we assume the following:

Assumption 4.4 $\widetilde{\rho}_i > 0 \ \forall i$.

Now the problem will not simply be to maximize expected readership, as each reader must be weighted by her "net value". The next proposition provides the Nash equilibrium condition in this case.

Proposition 4.3 *Suppose Assumptions 4.1-4.4 are satisfied and* $\lambda > 0$. *Then an equilibrium strategy profile* $\{s_\Gamma^*, s_\Theta^*\}$ *must satisfy* $s_\mu^\Gamma = s_\mu^\Theta, s_o^\Gamma = s_o^\Theta$ *and* $\frac{q'(s_\mu^*)}{q'(s_o^*)} = \frac{N_o \widetilde{\rho}_o \Phi_o}{N_\mu \widetilde{\rho}_\mu \Phi_\mu}$.

Proof. The profit equation for newspaper j can be expressed as

$$E(\Pi^j) = \widetilde{\rho}_\mu N_\mu H_\mu(\widehat{\Psi}_\mu) + \widetilde{\rho}_o N_o H_o(\widehat{\Psi}_o) - \overline{C}, \ j = \Gamma, \Theta.$$

The result follows from the first order conditions

$$\begin{aligned} \widetilde{\rho}_\mu N_\mu h_\mu(\widehat{\Psi}_\mu)\Phi_\mu q'(s_\mu^j) &= \vartheta \\ \widetilde{\rho}_o N_o h_o(\widehat{\Psi}_o)\Phi_o q'(s_o^j) &= \vartheta, \ j = \Gamma, \Theta. \end{aligned}$$

where ϑ is the Lagrange multiplier associated with this problem. Assumption 4.1 guarantees that the second order conditions are satisfied (see Lindbeck and Weibull, 1987).■

Proposition 4.3 tells us that now newspapers can discriminate across constituencies also on the basis of further information they may have. Other things equal, information supply will be higher in the constituency with larger $\widetilde{\rho}_i(\cdot) = \rho_i(\cdot) - \upsilon_i(\cdot)$. On the revenue side we can relate the readers' value for advertisers to factors as income, age, education etc. The net value of readers for newspapers will then take into account their location and be higher where readers are on average more valuable and lower where delivery costs are higher; we will use population density to capture this last element.

Finally, also the total size of constituencies, N_μ and N_o, (in terms of absolute population, or absolute electorate) should play a role. However, as we noticed at the start of this section, we can have both a "groups size effect" (like in Stromberg) as well as a "collective action effect" and we will approach empirical investigation with no prior about the sign of this variable.

We can therefore summarize our findings in the following testable proposition:

Theoretical Results *Other things equal, information supply is higher in constituencies with a closer electoral race, more densely populated, and where citizens are on average more valuable to advertisers. The effect of the size of electorate is uncertain.*

4.3 Evidence

We will proceed now to verify the compatibility of our theoretical results with data. Empirical investigation will concern the 1997 general election in the United Kingdom. For the purpose of this analysis we will use data on England, Scotland, and Wales only. As already noted in Chapter 2, Northern Ireland has been excluded as the political cleavages there are substantially different from the rest of the country. For a brief description of the UK political landscape and the 1997 election see section 5 in chapter 2.

4.3.1 The Data

Evidence provided is of two types. First, we will focus on information supply, using the electoral constituency as unit of observation. There were 641 such constituencies in England, Scotland, and Wales in 1997.

Three main sources of data will be used. First of all we need data about information supply by newspapers. For this purpose we will use a major national newspaper, "The Guardian". We will define information supply for each constituency as the number of articles that mention such constituency or one of its candidates during the last 30 days of the electoral campaign. This variable is indicated as $News$.

We will then use information about electoral results[4]. In particular, we will use this information to measure the marginality of a constituency. A first possibility is to focus on percentage differences and therefore use the following formula:

$$\frac{W - R}{W + R} \tag{4.3}$$

where W is the percentage of votes for the winning candidate and R the percentage for the runner up. The larger such indicator the lower the degree of marginality of the constituency. However, to capture the idea of marginality as the probability of casting a decisive vote, the absolute difference in votes between candidates might be a more appropriate indicator. We will consider both possibilities.

One problem with such indicators is that they measure election closeness *ex post*. A rational expectations assumption would work in favour of using such measures: in general, when using aggregate data, there is no reason to expect a systematic bias in expectations within a constituency. Nevertheless, voters' swings are not always well predicted by opinion polls, and this could generate non-random biases in voters' expectations[5].

One alternative possibility is to use past election results. Unfortunately this would also be quite problematic in our case. The main obstacle is that in between 1992 (year of the previous general election) and 1997 most constituency borders were changed, making therefore hard any comparison, even where the denominations were left unchanged. Moreover, in 1997 there were expectations of a large swing from the ruling party (Conservatives) to opposition parties (mainly the Labour): thus, previous election closeness could not represent a good measure of expected election closeness as this would crucially depend on who held the constituency.

For these reasons, to capture expected closeness we will use the BBC's "100 contested

[4]Boothroyd (2002).

[5]See Cox (1988) or Kunce (2001) for some problematic aspects of *ex post* indicators.

constituencies", the variable *bbc100* already described and used in chapter 2. This should capture information from polls and the general feeling about election closeness of BBC journalists.

We will also use information about the total number of electors in each constituency and the turnout percentage. With the first variable we try to test if the "group size effect" can actually prevail on the "collective action effect". The percentage of turnout indicates the extent of political participation (at least in the form of voting) and therefore can be broadly intended as a measure of interest in politics by the citizens of a given area.

Information about other possibly relevant characteristics of the constituency will be derived from the 1991 Census[6]. We include the population density, one of the key variables in our theoretical analysis, entering into the newspapers' cost function. Then we consider variables that can possibly give a representation of the social and economic conditions of the districts. Information on income is not available but proxies have been used, namely the unemployment rate and the percentage of citizenship with high qualifications (degree and higher). Age can have an influence on propensity to consume and consumption patterns (thus affecting how valuable a reader is to advertisers) and therefore has been included. Also, the percentage of inactive population (mainly retired, but also including students and permanently sick) is used. There are reasons (as well as anecdotic evidence[7]) to think that inactive population, in particular old or sick individuals, should be less valuable to advertisers, as they tend to consume less than average, or are less responsive to advertising.

One possible concern might derive from the fact that The Guardian, like most national newspapers in the U.K., is London based. This could bias the news in favour of London constituencies both because of a lower cost of news collection and, more generally, because of a larger sensitivity to a nearer environment. This could be particularly relevant for our results about population density, given that this variable is clearly higher in London than elsewhere. For this reason we include a Greater London control dummy, equal to 1

[6]The data we used were recorded at the level of districts, local administration entities with no direct link with electoral constituencies. Most constituencies are contained within the borders of a single district and these posed no problems. Others (around 25% of them) span over parts of different districts and in such cases data referred to districts have been weighted in order to get approximated constituency data. The weighting factors have been reconstructed by using the detailed description of constituencies (and their relations with districts and wards) contained in Rallings and Thrasher (1995).

[7]See Stromberg (2002).

for the Greater London constituencies.

Finally we also include a "big-shot" control. As some candidates have naturally a prominent position and bigger visibility during the electoral campaign, it seems necessary to be able to single out this effect from what we want to test. Therefore we introduce a dummy variable equal to 1 for constituencies where "big-shots" are candidates. By big-shot we intend all the candidates who have been ministers in any past government, the members of the current "shadow-cabinet", and the current leader of the Liberal Democratic Party[8].

We also provide evidence on citizens' use of newspapers across different constituencies. This helps us isolating the hypothesis that differentiated supply is actually a consequence of differentiated demand from the competing possibility that all citizens are interested in marginal constituencies. For this purpose we will use the 1997 British General Election Study, a post-election survey consisting of 3625 individual observations about people that were interviewed a short time after the election. Our sample will consists of 2807 observations. Among other questions, respondents were asked whether and how frequently they used to read newspapers during the electoral campaign, and which paper. In the U.K. the distinction between high quality and low quality (tabloid) papers is quite straightforward and commonly accepted. We can therefore separate regular users of quality papers during the electoral campaign (QP) from the rest of the population and try to assess the impact of marginality on the demand for political information. We also have information on the usage of local papers and this information can also be exploited to make our conclusions more robust. Data include a number of demographic and economic characteristics of the interviewed individuals, as well as a measure of ideological motivation.

All variables are described in more detail in the Appendix to this chapter and summary statistics are reported in table 4.1.

4.3.2 Empirical Specification

Using the dataset described above we intend to test the theoretical results reported at the end of section 4.2. Preliminary data analysis seems to suggest that a very limited number of constituencies get a disproportionate attention from media (see Tab. 4.2). For

[8]For further details see section 5 in chapter 2.

example almost 90% of constituencies have $News \leq 5$ while only 3 constituencies have $News > 100$. This suggests that the relationship we want to estimate could be highly non-linear. A linear regression would indeed deliver quite poor results. We will instead present estimates for the following equation:

$$\ln(News_i) = \alpha_0 + \alpha_1 D_i + \alpha_2' X_i + \alpha_3' Z_i + u_i, \; i = 1, ..., 641 \tag{4.4}$$

where:

D is a measure of the distance between candidates (winner and runner up) or a dummy for expected contested constituencies ($bbc100$), X is a three-dimensional vector of population density, size of the electorate, and turnout (therefore $\alpha_2' = [\alpha_{21}, \alpha_{22}, \alpha_{23}]$), and Z represents a set of control variables from the 1991 Census, plus the "big-shot" dummy ($\alpha_3' = [\alpha_{31}, \alpha_{32}, ...\alpha_{3k}]$). As usual u_i represents independent disturbance terms that have zero mean and are uncorrelated with the exogenous variables of the model. Estimation will be by OLS and will take into account potential heteroskedasticity. Specification tests will also be conducted, in particular a Box-Cox test to assess the log-linear functional form.

Almost all the parameters have an expected sign in terms of our model. However, our main parameter of interest is α_1. In general, we want to assess if α_1 is significantly different from zero. As discussed previously, we will consider several possible measures of constituencies' marginality.

The other variables serve as controls with respect to this aim; at the same time they are of interest for their own sake as we can use their estimates to assess the reliability of our model of the information market.

We will accomplish our task also by estimating newspaper readership at the individual level. The equation to be estimated in this case is given by

$$QP_i = \beta_0 + \beta_1 D_i + \beta_2' W_i + u_i, \; i = 1, ..., 2807 \tag{4.5}$$

where QP is a binary variable equal to 1 for a quality paper or a local paper reader, W is a vector of individual control variables including, among other covariates, income, education, sex and age.

4.3.3 Results

OLS estimates of equation 4.4 are reported in Table 4.3. We start by considering the role of election closeness. In column 1 and 2 we use *ex post* indicators of election closeness (based on percentage distance in column 1 and absolute distance in column 2). In both cases *ex post* distance has the expected sign and is significant at 5% level in column 1 and very close to it in column 2. When we use expected closeness (as captured by the dummy *bbc100* in column 3) the significance level increases substantially, reaching a nil p-value for the hypothesis of α_1 being zero. As discussed previously, the 1997 general election witnessed a large generalized shift of votes away from the Conservative party. This was to some extent expected and therefore the most interesting constituencies were the previously Conservative-held ones. In a sense, the final outcome was mainly decided in such constituencies and this should have increased the demand for information[9]. Thus, it is not surprising that when we focus on contested Conservative constituencies results get sharper. Actually, some constituencies may have been *ex post* very close just because the swing of votes has probably been larger than expected, making the Labour candidates winning (marginally) also in constituencies that never were marginal or Labour-held before. Therefore, *ex post* marginality could be an imperfect measure of expected salience.

In column 4 we also get rid of another *ex post* indicator, turnout, but results remain substantially unchanged. Thus, from this analysis we can safely conclude that expected marginality matters for information supply.

Other variables also show a high compatibility of our model with facts. Population density has the expected sign and is always significant at the 5% level, except in column 3, where it is significant at the 7.5% level.

The sign of other control variables also show good support for some of the Stromberg-type conclusions. In particular, and differently from Stromberg, we saw that the effect of group's magnitude is not necessarily uncontroversial. However, empirical evidence seems to suggest that the effect of the group size should overcome the potential collective action problem that size generates. In particular, in column 2 we use the absolute distance between candidates and therefore we isolate the potential "group size effect". However,

[9] In terms of our model, in general elections citizens care mainly about final policies: thus, marginality in one constituency is more relevant when it matters for the whole outcome of the election.

there are no noticeable differences between this and the other cases. Although the magnitude and significance of the electorate size are larger in column 2, these remain always positive and comfortably significant at the 5% or 1% level.

The same is not true for Turnout, that also should serve as a signal to newspapers about the degree of attention to political matters. The sign is always negative, although t-ratios are never satisfactory.

Good support also comes from the unemployment rate, that we use as a proxy for the level of well-being in a given constituency. On the other hand there is little evidence in favour of the relevance of other factors that the literature seems to have identified as determining information supply. Anecdotic evidence is reported of television programmes that have been suspended because watched mainly by the elderly, who where judged not valuable by advertisers. We find that constituencies with larger inactive population (mainly represented by retired) receive more attention from newspapers. It is clear that inactive people might have more time to devote to information gathering and when we come to election times retired people might also have all the incentives to put a disproportionate attention to political platforms.

Finally, average age and the percentage of people with high degrees do not seem to have significant effects, while there is some evidence of a "Greater London effect". A pure control variable is *big-shot*. Both the magnitude and the significance of *big-shot* are relevant but this does not come as a surprise nor is the consequence of any theoretical advance made in this paper.

In table 4.4 we turn to micro-level analysis and report probit estimates of equation 4.5. Most parameters show the expected sign, with education and income being overall the best explanatory variables. Sex, church attendance and length of residence in the area also show sizeable and significant effects. Our main variables of interest, however, are the measures of the size of the electorate and marginality. While for the significance of most other variables several explanations are possible, the electorate size and marginality have a strong relationship with voters' instrumental behaviour. Both come with the expected sign, quite independently of the marginality index used and the significance levels are rather reassuring. We can see this as further evidence that mass media behaviour during that electoral campaign was actually driven, at least partially, by instrumental demand for information rather than a broad and non-instrumental interest in the election.

4.4 Summary and conclusion

The purpose of this chapter has been to study the implications of instrumental voting behaviour for the political information market. This allows us to test instrumental voting theories by using data that have not been exploited so far for this purpose, namely data on information acquisition and mass media behaviour during electoral campaigns.

One central implication of instrumental voting is the positive linkage between election closeness and political participation. Both theoretical and empirical literature have mainly identified participation with electoral turnout. We focus instead on information acquisition and make more precise the idea that it should be higher when elections are expected to be closer. On the other side of the information market, profit maximizing mass media should therefore discriminate between different electoral constituencies according to their expected marginality. We do not observe newspapers' price discrimination in reality. However our research shows, both theoretically and empirically, that the media can have a different way to discriminate, namely targeting their attention (in terms of reported news) to marginal constituencies.

Moreover, research in communication studies and recent formal models, in particular Stromberg (2002), have pointed out that the media can be expected to target customers who are more valuable to advertisers, i.e. wealthier, better educated, younger. Our model gives an explicit empirical content to those predictions and, by using the electoral constituencies in the 1997 British election as units of observation, we can perform a formal empirical test of these conclusions. Evidence on Stromberg's conclusions is overall satisfactory: although not all our estimates are compatible with such results, we can safely conclude that there is enough evidence of newspapers targeting their news according to the electorate characteristics. We also provide empirical support for the idea that larger groups should receive more attention from the media, although we have shown that this conclusion does not necessarily follow from the theory.

Information supply can also be linked with the newspapers' cost function. We do not enter into the details of fixed costs, that represent a large part of the cost of producing a newspaper. However, we find that delivery costs could be relevant for the purpose of our analysis: in particular, information supply should be higher in more densely populated areas. This proposition too finds confirmation in our empirical investigation.

Our main purpose, however, was to show that information acquisition and news supply is, at least partially, driven by instrumental voting. We show that mass media, other things equal, tend to target marginal constituencies during electoral campaigns. This could be due to a genuine higher information demand arising in marginal constituencies as well as to a number of other reasons, like a general interest of the public in marginal constituencies, or the effort of party leaders to target marginal constituencies. To discriminate between these hypotheses we also provide evidence on voters' usage of newspapers and find that quality and local papers tend to be more demanded in marginal constituencies. Thus, our analysis seems to suggest a comfortable compatibility between instrumental voting behaviour and observed facts in the information market.

If we think that the media introduce a bias in the way people are informed about politics, something that has been left aside in the present work, and if this bias is in turn exploited by politicians, then we can speak of a "media-driven-bias" in public policy-making. In the context of our model this bias is combined with an "attention-bias" that should substantially drive politicians to target marginal constituencies.

This analysis is by no means conclusive and there are several margins for improvements and questions that further research should try to address.

On the theoretical side, the model of media competition is still quite simple. New insights could come from explicitly considering the advertising market and the possibility for newspapers to select the combination of political information, advertising and other news in the paper. Also, considering the possibility of entry and, more in general, different industry structures, could deliver interesting results as well as normative implications for regulating the media market[10]. Further research could help us understanding redistributive implications. In particular, and depending on the rules that regulate the relationship between central governments and local administration, we should expect marginal constituencies to benefit disproportionately of targetable benefits. This possibility deserves closer scrutiny in future research.

On the empirical side, improvements on our current knowledge are also possible. For what concerns our estimates, the relationship between news and closeness (as well as news and other variables) is clearly non-linear. We have chosen a log-linear specification and

[10] An analysis of this type with respect to politicians' accountability can be found in Besley and Prat (2002).

shown that it fits our data quite well. It would clearly be useful however to resort to non-parametric estimation. Above all there is the need for further data collection about both individuals and the media. More data about different newspapers (or other media), different elections and, possibly, different countries, could help us understand what is the robustness and the generality of our results and maybe to isolate the relevant institutional characteristics that induce differentiated behaviour.

4.5 Appendix: Description of Variables and Regression Results

Constituency level

- *News*. It is the number of articles appeared on the newspaper "The Guardian" during the last 30 days before the poll date and containing either a reference to the electoral constituency or the name of one of its candidates.
- $Dist\_P$ = percentage distance between the winning candidate and the runner-up in the 1997 election, given by the formula (W-R)/(W+R), where W = percentage of votes for the winning candidate, R = percentage of votes for the runner up.
- $Dist\_N$ = absolute distance between the winning candidate and the runner-up in the 1997 election, divided by 1000.
- $bbc100$ = dummy variable equal to 1 if the constituency has been included by the BBC among the "100 contested constituencies".
- *Density* = population density expressed as the number of residents per hectare divided by 1000.
- *Electorate* = total electorate in the constituency divided by 1000.
- $big-shot$ = dummy variable equal to 1 if one of the candidates in the constituency has been classified as a "big-shot". This means when one of the candidates is either a current or former minister, or a current member of the "shadow cabinet", or the leader of the Liberal-Democratic Party.
- *Unemployment%* = unemployment rate, expressed as total unemployed over active population, multiplied by 100.
- *Inactive%*= percentage of inactive population. This is the total of retired, students, permanently sick and other inactive over total residents multiplied by 100.
- *Age* = average age.
- *HighD* = percentage of residents with high qualifications, defined as the number of residents with degree or higher title over the total residents, multiplied by 100;

- *GLondon* = dummy variable equal to 1 for the greater London constituencies.

Individual level

- *QP* = dummy variable equal to 1 if the respondent is a regular reader of The Times, The Guardian, The Independent, The Financial Times, The Daily Telegraph, or any local newspaper.
- *Education* : respondent's education level. Categorical variable from 1 to 7.
- *Income*: total household income from all sources before tax. Categorical variable from 1 to 16.
- *Age* : respondent's age (>18).
- *Sex*: 1=male
- *Married*: 1=yes (=1 also if "living as married")
- *Asian*: =1 if Indian, Pakistani, Bangladeshi, Chinese, Other Asian.
- *Black*: =1 if Black African, Black Caribbean, Other Black.
- *Churchgoer*: "Apart from such special occasions as weddings, funerals and baptisms and so on, how often do you attend services or meeting connected with your religion?". Categorical variable from 1 (never or practically never) to 8 (once a week or more).
- *Lenght of Residence* : Answer to the question: "How long have you lived in this neighbourhood?".
- *Ideology*: derived from individual placement on a left (0) to right (10) scale. Ideology=0 if left-right=5, Ideology=1 if left-right=4 or 6 etc.
- *Registered*: 1=yes.
- *Voted92*: =1 if voted in 1992 general election (self reported).
- *Economic Activity*. Categorical variable:

1. "in paid work for at least 10 hours in week" or "waiting to take up paid work already accepted"; 1498 obs.;

2. "in full time education (not paid for by the employer, including on vacation". 9 obs.;

3. "on government training/employment programme". 64 obs.;

4. "unemployed". 127 obs.;

5. "permanently sick or disabled". 131 obs.;

6. "wholly retired from work". 642 obs.;

7. "looking after the home". 324 obs.;

8. "other". 18 obs.

Table 4.1: Summary statistics

	Obs	Mean	Std. Dev.	Min	Max
News	641	3.8658	18.4735	0 (app. to 0.01)	388
Log(News)	641	-1.3847	2.8319	-4.6052	5.961
Dist_P	641	0.2972	0.2006	0.0012	0.8219
Dist_N	641	10859	6906	53	33759
BBC100	641	0.156	0.3631	0	1
Density	641	1.6524	1.9092	0.0088	11.6298
Electorate/1000	641	66.5437	8.0574	22.983	101.68
Turnout%	641	71.3165	5.6359	51.4	82.2
BS	641	0.078	0.2684	0	1
Age	641	37.8866	1.881	32.8793	46.533
Inactive%	641	51.32	3.2413	38.858	61.0908
Unemployment%	641	9.45	3.8	2.868	22.4896
HighD%	641	7.0963	3.743	1.4891	25.084
GLondon	641	0.1154	0.3198	0	1
QP	2807	0.1778	0.3824	0	1
Education	2807	3.6021	2.1637	1	7
Income	2807	7.0495	4.587	1	16
Age	2807	48.3035	17.517	18	94
Sex	2807	0.4653	0.4989	0	1
Married	2807	0.5885	0.4922	0	1
Asian	2807	0.0185	0.1349	0	1
Black	2807	0.0089	0.094	0	1
Churchgoer	2807	1.9882	2.6079	0	7
Length of Residence	2807	19.5248	17.9378	0	94
Ideology	2807	1.9291	1.7758	0	6
Registered	2807	0.9865	0.1156	0	1
Voted92	2807	0.7973	0.4021	0	1

Table 4.2: The variable "News"

News	Frequency	Percent	Cumulate
0	265	41.34	41,34
1	165	25.74	67.08
2	74	11.54	78.63
3	28	4.37	83
4	18	2.81	85.80
5	24	3.74	89.55
6-10	25	3.9	93.45
11-20	19	2.96	96.41
21-30	10	1.56	97.97
31-40	3	047	98.44
41-50	3	047	98.91
51-100	4	0.62	99.53
>100	3	0.47	100

Table 4.3: Information Supply (OLS)

Dependent Variable = Log(News)

	1	2	3	4
Dist_P	-1.4808 (-1.960)			
Dist_Nx1000		-0.0366 (-1.924)		
BBC100			1.1889 (4.306)	1.0415 (3.84)
Density	0.2448 (2.076)	0.2521 (2.153)	0.2081 (1.783)	0.2535 (2.227)
Electorate/1000	0.0321 (2.311)	0.3891 (2.884)	0.0334 (2.463)	0.336 (2.46)
Turnout	-0.04 (-1.406)	-0.0305 (-1.138)	-0.0472 (-1.801)	
Big shot	3.4017 (8.815)	3.3891 (8.726)	3.4925 (8.955)	3.4918 (8.899)
Age	-0.1074 (-1.268)	-0.1071 (-1.264)	-0.0701 (-0.83)	-0.07 (-0.836)
Inactive	0.1449 (2.022)	0.1423 (1.989)	0.1257 (1.753)	0.1209 (1.704)
Unemployment	-0.1789 (-2.544)	-0.179 (-2.555)	-0.1985 (-2.847)	-0.1625 (-2.414)
HighD	0.0327 (0.857)	0.0284 (0.742)	0.039 (1.036)	0.0322 (0.862)
GLondon	0.7046 (1.589)	0.7094 (1.605)	0.8574 (1.971)	0.8016 (1.855)
Constant	-2.8828 (-0.764)	-3.919 (-1.082)	-3.3395 (-0.948)	-6.8138 (-2.368)
Obs	641	641	641	641
R-squared	0.1662	0.166	0.1818	0.1779

Note: robust standard errors. T-statistics in parenthesis

Table 4.4: Information Demand
(Probit marginal effects)

Dependent Variable = QP

	1	2	3
Education	0.0368 (10.20)	0.0368 (10.03)	0.0367 (9.99)
Income	0.0178 (9.12)	0.0184 (9.30)	0.0185 (9.31)
Age	0.0041 (1.67)	0.0045 (1.80)	0.0045 (1.80)
Age2	0.0002 (0.07)	-0.00002 (0.01)	-0.00004 (0.02)
Sex	0.0584 (4.14)	0.0605 (4.23)	0.0604 (4.22)
Married	-0.0037 (0.24)	-0.0045 (0.29)	-0.0045 (0.30)
Asian	0.0669 (1.12)	0.0545 (0.93)	0.0538 (0.92)
Black	0.1124 (1.36)	0.0734 (0.95)	0.0717 (0.93)
Churchgoer	0.0107 (4.36)	0.0114 (4.60)	0.0115 (4.62)
Length of Residence	-0.0009 (2.20)	-0.0011 (2.46)	-0.0011 (2.47)
Registered	-0.0969 (1.45)	-0.3821 (1.50)	-0.3955 (1.55)
Voted92	-0.0104 (0.58)	-0.0112 (0.61)	-0.0111 (0.60)
Ideology	0.0199 (5.41)	0.0196 (5.28)	0.0196 (5.28)
GLondon	0.0404 (1.64)	0.0450 (1.78)	0.0437 (1.74)
Electorate	-0.0029 (3.98)	-0.0021 (2.92)	-0.002 (2.85)
Dist_P	-0.1785 (4.87)		
BBC100		0.0474 (2.24)	0.0453 (2.14)
Big shot			-0.0225 (0.85)
Economic Activity	yes	yes	yes
Obs.	2807	2807	2807
Log-Likelihood	-1053.1073	-1061.8214	-1061.4073
Pseudo-R2	0.1983	0.1917	0.192

Note: robust standard errors. T-statistics in parenthesis

Chapter 5

Conclusion

Information plays a key role in decision-making. This consideration has delivered a completely new perspective in economic theory, where the elegance of the Arrow-Debreu general equilibrium model has often been replaced by partial equilibrium analysis dealing with specific aspects of the economy. Decisions taken in the political market of democratic systems make no exception, and crucially rely on information availability and its quality. For a number of different reasons, good information is important for policy-makers, for candidates, for lobbies, and for the common citizen.

This work tries to contribute to a better understanding of the role of information in elections, its impact on voting behaviour and, ultimately, on public policy. It is argued that to learn about the consequences of information we need to start from its determinants. Not all actors in the political arena have the same incentives to be informed. While professional politicians have sufficient motivations to gather information on political and economic matters, the same is not true for common citizens, whose impact on public decision-making is often extremely limited. Thus, political information acquisition by simple voters, as many other social phenomena involving collective action problems, remains largely unexplained and constitutes a still puzzling phenomenon to social scientists.

From this apparent *empasse*, it is possible to make substantial progresses when we recognize that the economic method consists not much in *explaining* behaviour in a strict sense as instead in linking *changes* in observed behaviour to *changes* in observable constraints. This is the approach taken in this thesis, where the consideration of individual

motivations has been coupled with a focus on opportunity and ability to learn: although the motivation to be informed could well derive mainly from individual unobservable tastes, the opportunity and the ability to learn will eventually leave their mark on the amount and the type of political knowledge that citizens possess. If information has an impact on public policy-making, then this positive analysis can also deliver a number of normative implications. Thus, finding causal relationships between heterogeneity in political knowledge and observable constraints permits to evaluate different institutional and market arrangements and opens the possibility of actual policy recommendations.

The central idea developed in this thesis is that exogenously including imperfect information in electoral models is not enough. Individual incentives to be informed are correlated with other incentives to participate in public life: this makes the effects of information hardly identifiable in empirical research. A theory of information acquisition is therefore needed to isolate the variables that affect political knowledge from those determining political participation in general. By modelling information acquisition as an individual production we can identify the impact of political knowledge on electoral turnout and conclude that being informed on political matters significantly increases the likelihood of voting. It is important, however, to distinguish the flow of information received during an electoral campaign from the prior stock of political knowledge. Such knowledge is the outcome of each individual's history, from parents' influence to direct personal experiences: thus, prior perception of political matters is generally very diverse across the population. We show that such diversity also plays an important role when coming to information acquisition and the impact of political knowledge on turnout.

This study shows that availability of news on mass media and individual resources to acquire, process, and retain information are good explanatory variables of voters' political knowledge. As a growing research literature stresses the importance of information for agents' (in this case public officials) accountability, we can conclude that mass media and voters' personal resources play a crucial role in democratic decision-making. Formally democratic institutions are emptied of their substantial content if political information is unavailable or beyond most voters' reach.

Moreover, resources and costs that affect information acquisition are asymmetrically distributed across the population: this introduces the possibility that political entrepreneurs, mass media and possibly other actors in the electoral game will influence the final

outcome by discriminating among voters. Politicians can target more responsive segments of the population, thus favouring those who have larger opportunity to be informed. In this work we show that this asymmetry could result in an higher political weight given to the rich and in public policy systematically biased in their favour. Considering this asymmetry casts serious doubts on the ability of the standard Downsian framework to predict redistributive policy. In particular we show the weakness of the link between inequality and redistribution that can be derived as a corollary of the median voter theorem. Inequality may well increase the median voter' desire for redistribution; however, at the same time it also increases inequality in political awareness, thus rendering the poor less capable of pursuing their interest on the political market.

Coming to information supply, we provide evidence of mass media targeting voters who demand more information or are more valuable to advertisers: this introduces a further bias in the distribution of information. Various groups and organizations may actually have similar incentives and can therefore be expected to exploit any informational advantage they might have for their own purposes. If such information biases are transferred into a policy bias, as we argued in this thesis, then mass media can leave their own mark on public policy by simply maximizing their profits, i.e. excluding any possible bias introduced for ideological or "capture" reasons. We also provide new evidence of voters' instrumental behaviour by showing that election closeness matters for information acquisition, as one would expect if political information is, among other reasons, acquired to make better electoral choices.

Thus, summarizing, this thesis provides theoretical arguments and empirical evidence of the importance of information for voting behaviour, the role of personal and environmental constraints in determining political knowledge, and the active role of mass media in determining political awareness. We also provide theoretical reasoning unveiling the potential impact of political information on democratically determined public policy.

The findings of this study do not pretend to be conclusive and further research is necessary in order to evaluate the robustness of our results when different institutional arrangements are considered. A number of questions deserve further investigation. How electoral rules affect information acquisition and information transmission during electoral campaigns? What should we expect under proportional representation? What is the impact of state-owned media on the information market? Such questions are not just

a theoretical curiosity: they are somehow central in many countries' current debates on electoral reform and mass media regulation. Also, we only partially addressed the link between information and public policy. Evidence on policy-makers targeting of informed groups would make us more sanguine about the practical relevance of many of the findings of this thesis, from the unbalanced representation of different income groups to the information bias introduced by mass media. It would certainly be useful, for example, to test whether marginal constituencies are actually targeted by policy-makers and if information has a role in this targeting.

Certainly the possibility to answer these and many more questions depends on advances in a number of other grounds. Theoretical research is slowly incorporating imperfect information in electoral models, unveiling a number of previously ignored effects of electoral competition. These can be particularly interesting when coupled with the insights offered by new models of electoral competition that consider a multidimensional policy space. At the same time, this trend in theoretical literature is only beginning to affect empirical research. A large number of questions is still waiting to receive the attention it deserves. This will hopefully stimulate new data collection on information and elections. It seems fair to say that at the moment our knowledge of some phenomena is severely constrained by limited availability of appropriate data. One of the main tasks of future research should be to close this gap.

Bibliography

[1] Abramson, P.R., Silver, B.D., and Anderson, B.A. (1987): The effects of question order in attitude surveys: the case of the SRC/CPS citizen duty items, *American Journal of Political Science,* **31,** 900-908.

[2] Aldrich, J.H. (1993): Rational choice and turnout, *American Journal of Political Science,* **37,** 246-278.

[3] Alesina, A. (1987): Macroeconomic policy in a two-party system as a repeated game, *Quarterly Journal of Economics, **102**, 651-78.*

[4] Alesina, A. (1988): Credibility and policy convergence in a two-party system with rational voters, *American Economic Review,* **78**, 796-806.

[5] Alesina, A. and Cukierman, A. (1990): The politics of ambiguity, *Quarterly Journal of Economics*, **105** (4), 829-850.

[6] Alesina, A. and Rodrick, D. (1994): Redistributive politics and economic growth, *Quarterly Journal of Economics*, **109,** 465-490.

[7] Aristotle (1946): *The Politics of Aristotle.* Trans. by E. Barker. Oxford: Oxford University Press.

[8] Arrow, K.J. (1951): *Social choice and individual values*, New York: Wiley.

[9] Arrow, K.J. (1986): The demand for information and the distribution of income, *Institute for Mathematical Studies in the Social Sciences, Stanford University, Technical Report n. 494.*

[10] Austen-Smith, D and Banks, J. (1996): Information aggregation, rationality, and the Condorcet jury theorem. *American Political Science Review,* **90**, 34-45.

[11] Banks, J.S. and Sundaram, R.K. (1998): Optimal retention in agency problems, *Journal of Economic Theory*, **82**, 293-323.

[12] Barro, R.J. (1973): The control of politicians: an economic model, *Public Choice,* **14**, 19-42.

[13] Bartels, L.M. (1988): *Presidential primaries and the dynamics of public choice.* Princeton: Princeton University Press.

[14] Bartels, L.M. (1993): Messages received: the political impact of media exposure, *American Political Science Review,* **87,** 267-284.

[15] Bartels, L.M. (1996): Uninformed voters: information effects in presidential elections, *American Journal of Political Science*, **40**, 194-230.

[16] Benabou, R. (2000): Unequal societies: income distribution and the social contract, *American Economic Review, 90, 96-129.*

[17] Berelson, B.R., Lazarsfeld, P.F., and McPhee, W.N. (1954): *Voting.* Chicago: The University of Chicago Press.

[18] Berg, S. (1993): Condorcet's jury theorem, dependency among jurors, *Social Choice and Welfare*, **10**, 87-96.

[19] Bertola, G. (1993): Factor shares and savings in endogenous growth, *American Economic Review*, **83**, 1184-1198.

[20] Besley, T. and Burgess, R. (2002): The political economy of government responsiveness: theory and evidence from India, forthcoming, *Quarterly Journal of Economics.*

[21] Besley, T. and Case, A. (1995a): Incumbent behaviour, vote-seeking, tax-setting, and yardstick competition, *American Economic Review*, **85**, 25-45.

[22] Besley, T. and Case, A. (1995b): Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits, *Quarterly Journal of Economics*, **110**, 769-798.

[23] Besley, T. and Case, A. (2002): Political institutions and policy choices: evidence from the United States, *mimeo*, LSE and Princeton.

[24] Besley, T. and Coate, S. (1997): An economic model of representative democracy, *Quarterly Journal of Economics*, **108**, 85-114.

[25] Besley, T. and Coate, S. (1998): Sources of inefficiencies in a representative democracy: a dynamic analysis, *American Economic Review*, **88**, 139-156.

[26] Besley, T. and Pande, R. (1998): Read my lips: the political economy of information transmission, *IFS* Working Paper, 155.

[27] Besley, T. and Prat, A. (2001): Handcuffs for the grabbing hand? Media capture and government accountability, *mimeo*, LSE.

[28] Bishop, G.F., Oldendick, R.W., and Tuchfarber, A.J. (1984): Interest in political campaigns: the influence of question order and electoral context, *Political Behaviour,* **6**, 159-169.

[29] Black, D. (1948): The decisions of a committee using a special majority, *Econometrica*, **16**, 245-261.

[30] Blackwell, D. (1951): The comparison of experiments, in *Proceedings, Second Berkeley Symposium on Mathematical Statistics and Probability, 93-102.* University of California Press.

[31] Blumler, J. and McQuail, D. (1968): *Television in politics: its uses and influences.* London, Faber.

[32] Bollen, K.A., Guilkey, D.K., and Mroz, T.A. (1995) : "Binary outcomes and endogenous explanatory variables: tests and solutions with an applications to the demand for contraceptive use in Tunisia", *Demography*, **32**, 111-131.

[33] Boothroyd, D. (2002): United Kingdom Election Results, *http://www.election.demon.co.uk/.*

[34] Brady, H.E. and Sniderman, P.M. (1985): Attitude attribution: a groups basis for political reasoning, *American Political Science Review*, **79**, 1061-78.

[35] Caillaud, B. and Tirole, J. (1997): Voters homogeneity and aggregation of information, mimeo, CERAS-ENPC, Paris.

[36] Calvert, R.L. (1985): Robustness of the multidimensional voting model: candidate motivations, uncertainty and convergence, *American Journal of Political Science*, **29**, 69-95.

[37] Calvert, R.L. (1986): Models of imperfect information in politics, vol. 6 of *Fundamentals of Pure and Applied Economics,* eds. J. Lesourne and H. Sonneschein. Harwood Academic Publishers.

[38] Campbell, J.E., Converse, P.E., Miller, W.E., and Stokes, D.E. (1960): *The American voter.* New York: John Wiley.

[39] Coate, S. and Morris, M. (1995): On the form of transfers to special interests groups, *Journal of Political Economy*, **103**, 1210-1235.

[40] Condorcet, Marquis de (1785): *Essai sur l'application de l'analyse a la probabilite des decisions rendues a la probabilite des voix.* Paris: De l'imprimerie royale.

[41] Converse, P.E. (1964): The nature of belief systems in mass public. In *Ideology and Discontent,* ed. D.E. Apter, New York: Free Press.

[42] Coughlin, P.J. (1992): *Probabilistic voting theory*, Cambridge University Press.

[43] Coughlin, P. and Nitzan, S. (1981): Directional and Local Electoral Equilibria with Probabilistic Voting, *Journal of Economic Theory*, **24**, 226-39.

[44] Cox, G. (1988): Closeness and turnout: a methodological note, *The Journal of Politics*, **50/3**, 768-775.

[45] Cox, G.W. and M.C. Munger (1989): "Closeness, expenditure, turnout: the 1982 U.S. House elections", *American Political Science Review,* **83** (1),217-32.

[46] Crawford, V.P. and Sobel, J. (1982): Strategic information transmission, *Econometrica*, **50**, 1431-1451.

[47] Delli Carpini, M.X. and Keeter, S. (1996): *What Americans know about politics and why it matters.* Yale University Press.

[48] Djankov, S., McLiesh, C., Nenova, T. and Shleifer, A. (2001): Who owns the media?, paper presented at the STICERD workshop on *Mass Media Economics,* June 2002.

[49] Downs, A. (1957): *An Economic Theory of democracy.* New York, Harper & Row.

[50] Feddersen, T.J. and Pesendorfer, W. (1996): The swing's voter's curse, *American Economic Review,* **86**, 408-424.

[51] Feddersen, T.J. and Pesendorfer, W. (1997): Voting behavior and information aggregation in elections with private information, *Econometrica,* **65**, 1029-1058.

[52] Feddersen, T.J. and Pesendorfer, W. (1998): Convicting the innocent: the inferiority of unanimous jury verdicts, *American Political Science Review,* **92**, 23-35.

[53] Feddersen, T.J. and Pesendorfer, W. (1999): Abstention in elections with asymmetric information and diverse preferences, *American Political Science Review,* **93**, 381-398.

[54] Ferejohn, J. (1986): Incumbent performance and electoral control, *Public Choice,* **50**, 5-26.

[55] Ferejohn, J. and Kuklinski, J. (eds.) (1990): *Information and democratic processes.* Urbana & Chicago, University of Illinois Press.

[56] Finkel, S.E. (1993): Reexamining the 'Minimal Effects' model in recent presidential campaigns, *Journal of Politics,* **55**, 1-21.

[57] Foley, D. (1967): Resource allocation and the public sector, *Yale Economic Essays,* **7**, 45-98.

[58] Foster, C. (1984): The performance of rational voter models in recent Presidential elections, *American Political Science Review,* **78**, 678-690.

[59] Franklin, C.H. (1991): Eschewing obfuscation? Campaigns and the perceptions of U.S. Senate incumbents, *American Political Science Review,* **85**, 1193-1214.

[60] Gans, J.S. and Smart, M. (1996): Majority voting with single-crossing preferences, *Journal of Public Economics,* **59**, 219-237.

[61] Grandmont, J.M. (1978): Intermediate preferences and the majority rule, *Econometrica,* **46**, 317-330.

[62] Grofman, B., Collet, C., and Griffin, R. (1998): Analyzing the turnout-competition link with aggregate cross-sectional data, *Public Choice*, **95**, 233-246.

[63] Grofman, B. and Norrander, B. (1990): Efficient use of reference group cues in a single dimension, *Public Choice*, **64**, 213-217.

[64] Grossman, G. M. and Helpman, E. (1999): Competing for endorsements, *American Economic Review*, **89**, 501-24.

[65] Guilkey, D.K., Mroz, T.A., and Taylor, L. (1992): Estimation and testing in simultaneous equations models with discrete outcomes using cross section data, unpublished manuscript.

[66] Harrington, J.E. (1993): Economic policy, economic performance, and elections, *American economic review*, **83**, 27-42.

[67] Heath, A., Jowell, R., Curtice, J.K., and Norris, P. (1999): "The British General Election Study, 1997", available at http://www.data-archive.ac.uk/.

[68] Hicks, A.M. and Swank, D.H. (1992): Politics, institutions, and welfare spending in industrialized democracies, 1960-1982, *American Political Science Review*, **86**, 658-674.

[69] Hill, K.Q. and Leighley, J.E. (1992): The policy consequences of class bias in State electorates, *American Journal of Political Science*, **36**, 351-365.

[70] Hill, K.Q., Leighley, J.E., and Hinton-Andersson, A. (1995): Lower class mobilization and policy linkage in the US States, *American Journal of Political Science*, **39**, 75-86.

[71] Hinich, M.J. (1977): Equilibrium in spatial voting: the median voter result is an artifact, *Journal of Economic Theory*, **16**, 208-219.

[72] Hotelling, H. (1929): Stability in competition, *Economic Journal*, **39**, 41-57.

[73] Husted, T.A. and Kenny, L.W. (1997): The effect of the expansion of the voting franchise on the size of government, *Journal of Political Economy*, **105**, 54-82.

[74] Iyengar, S. and Kinder, D.R. (1987): *News that matters.* Chicago: The University of Chicago Press.

[75] Key, Jr., V. (1966): *The responsible electorate: rationality in Presidential voting 1936-1960.* New York: Vintage.

[76] Kihlstrom, R. (1974): A general theory of demand for information about product quality, *Journal of Economic Theory,* **8,** 413-439.

[77] Kramer, G.H. (1972): Sophisticated voting over multidimensional choice spaces, *Journal of Mathematical Sociology,* **2**, 165-180.

[78] Kunce, M. (2001): Pre-election polling and the rational voter: evidence from state panel data (1986-1998), *Public Choice*, **107**, 21-34.

[79] Lambert, P.J. (1993): *The distribution and redistribution of income: a mathematical analysis.* Manchester University Press.

[80] Lau, R. and Sears, D. (eds.) (1986): *Political cognition.* Hillsdale, Erlbaum.

[81] Lazarsfeld, P., Berelson, B., and Gaudet, H. (1944): *The people's choice.* New York: Columbia University Press.

[82] Ledyard, J.O. (1984): The pure theory of large two-candidates elections, *Public Choice*, **44**, 7-41.

[83] Leighley, J.E. and Nagler, J. (1992): Individual and systemic influences on turnout: who votes? 1984. The *Journal of Politics*, **54**, 718-740.

[84] Lijphart, A. (1997): Unequal participation: democracy's unresolved dilemma, *American Political Science Review*, **91**, 1-14.

[85] Lindbeck, A. and Weibull, J.W. (1987): Balanced-budget redistribution as the outcome of political competition, *Public Choice*, **52**, 273-97.

[86] Lindert, P. (1996): What limits social spending?, *Explorations in Economic History*, **33**, 1-34.

[87] Lippmann, W. (1922): *Public opinion*, New York: Free Press.

[88] Lohmann, S. (1998): An information rationale for the power of special interests, *Americal Political Science Review,* **92**, 809-27.

[89] Lupia, A. (1992): Busy voters, agenda control, and the power of information. *American Political Science Review,* **86**, 390-404.

[90] Lupia, A. (1994): Shortcuts versus encyclopedias: information and voting behaviour in California insurance reform elections, *American Political Science Review,* **88**, 63-76.

[91] Lupia, A. and McCubbins, M.D. (1998): *The democratic dilemma: can citizens learn what they need to know?* New York: Cambridge University Press.

[92] Margolis, H. (2001): Pivotal voting, *Journal of Theoretical Politics,* **13**, 111-116.

[93] Margolis, H. (2002): Pivotal voting and the Emperor's new clothes. *Social Choice and Welfare,* **19**, 95-111.

[94] Matsusaka, J.G. (1995): Explaining voter turnout patterns: an information theory, *Public Choice,* **84,** 91-117.

[95] Matsusaka, J.G. and Palda, F. (1999): Voter turnout: how much can we explain?, *Public Choice,* **98,** *431-446.*

[96] Meltzer, A. and Richards, S. (1981): A rational theory of the size of government, *Journal of Political Economy,* **89**, 914-927.

[97] McCombs, M. and Shaw, D. (1972): The agenda setting function of mass media, Public Opinion Quarterly, **36**, 176-187.

[98] McKelvey, R.D. (1986): Covering, dominance, and institution-free properties of social choice, *American Journal of Political Science,* **30**, 283-314.

[99] McKelvey, R.D. and Ordeshook, P.C. (1984): Rational expectations in elections: some experimental results based on a multidimensional model, *Public Choice,* **44**, 61-102.

[100] Mill, J.S. (1835) [1963]: De Toqueville on democracy in America, in *Essays on politics and culture,* Himmelfarb, G. (ed.). Doubleday, New York

[101] Mill, J.S. (1861): [1972] *Representative government.* Everyman Edition, London.

[102] Miller, N.R. (1980): A new solution set for tournaments and majority voting, *American Journal of Political Science*, **24**, 68-96.

[103] Mondak, J. (1994): Cognitive heuristics, heuristic processing, and efficiency in political decision making. In *New Directions in Political Psychology,* ed. Delli Carpini, M.X., Huddy, L.,Shapiro, R.Y. Greenwich-Conn.: JAI Press.

[104] Morton, R.B. (1991): Groups in rational turnout models, *American Journal of Political Science*, **35**, 758-776.

[105] Neuman, W.R. (1986): *The paradox of mass politics: knowledge and opinion in the American electorate.* Cambridge: Harvard University Press.

[106] Nie, N.H., Verba, S., and Petrocik, J.R. (1979): *The changing American voter.* Cambridge: Harvard University Press.

[107] Noam, E.M (1987): A public and private choice model of broadcasting, *Public Choice*, **55**, 163-187.

[108] Norris, P., Curtice, J., Sanders, D., Scammell, L., and Semetko, H.A. (1999): *On message: communicating the campaign.* London: Sage.

[109] Osborne, M.J. and Slivinski, A. (1996): A model of political competition with citizen-candidates, *Quarterly Journal of Economics*, **111**, 65-96.

[110] Palfrey, T.R. and Poole, K.T. (1987): The relationship between information, ideology, and voting behavior, *American Journal of Political Science,* **31**, 511-530.

[111] Palfrey, T.R. and Rosenthal, H. (1985): Voter participation and strategic uncertainty, *American Political Science Review*, **79**, 62-78.

[112] Pateman, C. (1970): *Participation and democratic theory.* Cambridge: Cambridge University Press.

[113] Patterson, S.C. and Caldeira, G.A. (1983): Getting out the vote: participation in Gubernatorial elections, *American Political Science Review*, **77**, 675-689.

[114] Perotti, R (1994): Income distribution and investment, *European Economic Review*, **38**, 827-835.

[115] Perotti, R. (1996): Growth, income distribution and democracy: what the data say, *Journal of Economic Growth*, **1**, 149-187.

[116] Persson, T. and Tabellini, G. (1994): Is inequality harmful for growth?, *American Economic Review*, **84**, 600-621.

[117] Peterson, P.E. and Rom, M. (1989): American federalism, welfare policies, and residential choices, *American Political Science Review*, **83**, 711-728.

[118] Plott, C.R. (1967): A notion of equilibrium and its possibility under majority rule, *American Economic Review*, **57**, 787-806.

[119] Pomper, G.M. (1972): From confusion to clarity: issues and American voters, 1956-1968. *American Political Science Review, **66**, 415-428.*

[120] Popkin, S.L. (1991): *The reasoning voter.* Chicago: The University of Chicago Press.

[121] Radner, R. and Stiglitz, J.E. (1984): A nonconcavity in the value of information, in *Bayesian models in economic theory,* eds. M. Boyer and R.E.Kihlstrom. North Holland.

[122] Rallings, C. and Thrasher M. (eds.) (1995): *Media guide to the new parliamentary constituencies*, Local Government Chronicle Elections Centre for BBC, ITN, PA News and Sky.

[123] Riker, W.H. and Ordeshook, P.C. (1968): A theory of the calculus of voting, *American Political Science Review,* **62,** 25-42.

[124] Roberts, K. (1977): Voting over income tax schedules, *Journal of Public Economics*, **8**, 329-340.

[125] Roemer, J.E. (1999): The democratic political economy of progressive income taxation, *Econometrica*, **67**, 1-19.

[126] Rogoff, K. (1990): Equilibrium political budget cycles, *American Economic Review*, **80**, 21-36.

[127] Rogoff, K. and Sibert, A. (1988): Elections and macroeconomic policy cycles, *Review of Economic Studies*, **55**, 1-16.

[128] Romer, T. (1975): Individual welfare, majority voting, and the properties of a linear income tax, *Journal of Public Economics*, **4**, 163-185.

[129] Rousseau, J.-J. (1762) [1968]: *The social contract*, Harmondsworth, Eng. Penguin.

[130] Saint Paul, G. (1994): The dynamics of exclusion and fiscal conservatism", CEPR Discussion Paper **998**.

[131] Saint Paul, G. and Verdier, T. (1997): Power, distributive conflicts, and multiple growth paths, *Journal of Economic Growth*, **2**, 155-168.

[132] Sanders, M.S. (2001): Uncertainty and turnout, *Political Analysis*, **9**, 45-57.

[133] Sartori, G. (1987): *The theory of democracy revisited*, Chatham, NJ: Chatham House Publishers.

[134] Schumpeter, J.A. (1942): *Capitalism, socialism and democracy.* New York, Harper and Row.

[135] Sen, A. (1981): *Poverty and famines*, Oxford, Oxford University Press.

[136] Sen, A. (1984): Food battles: conflicts in the access to food, *Food and Nutrition*, **10**, 81-89.

[137] Shachar, R. and Nalebuff, B. (1999): Follow the leader: theory and evidence on political participation, *American Economic Review,* **89**, 525-547.

[138] Shannon, C.E. (1948): A mathematical theory of communication, *The Bell System Technical Journal*, **27**, 379-423.

[139] Shepsle, K.A. and Weingast, B.R. (1984): Uncovered sets and sophisticated voting outcome with implications for agenda institutions, *American Journal of Political Science*, **28**, 49-74.

[140] Sniderman, P.M., Brody, R.A., and Tetlock, P.E. (1991): *Reasoning and choice: explorations in political psychology.* New York: Cambridge University Press.

[141] Sobel J. (1985): A theory of credibility, *Review of Economic Studies*, **52**, 557-573.

[142] Spence,M. and Owen, B. (1977): Television programming, monopolistic competition and welfare, *Quarterly Journal of Economics*, **91**, 103-26.

[143] Stromberg, D. (2001): Radio's impact on New Deal spending, mimeo, Princeton University.

[144] Stromberg, D. (2002): Mass media competition, political competition and public policy, forthcoming, *Review of Economic Studies.*

[145] Uhlaner, C.J. and Grofman, G. (1986): The race may be close but my horse is going to win: wish fulfillment in the 1980 presidential election, *Political Behavior*, **8**, 101-129.

[146] Verba, S. (1993): The uses of survey research in the study of comparative politics: issues and strategies, *Historical Social Research*, **18**, 55-103.

[147] Verrecchia, R.E. (1980): Information acquisition in a noisy rational expectations economy. *Econometrica,* **50**, 1415-1430.

[148] Wittman, D.A. (1995): *The myth of democratic failure: why political institutions are efficient.* Chicago: University of Chicago Press.

[149] Wolfinger, R.E. and Rosenstone, S.J. (1980): *Who votes?.* New Haven: Yale University Press.

[150] Young, P. (1988): Condorcet's theory of voting, *Americal Political Science Review,* **82,** 1231-44.

[151] Zaller, J. (1989): Bringing Converse back in: modeling information flow in political campaigns, *Political Analysis*, **1**, 181-234.

[152] Zaller, J. (1992): *The nature and origins of mass opinion.* New York, Cambridge University Press.

[153] Zaller, J. (1995): The myth of massive media effects revived: empirical support for a discredited idea, in *Political persuation and attitude change,* ed. Mutz, D., Sniderman, P.M., and Brody, R.A. Ann Arbor: University of Michigan Press.